

193061

P- 194

THE FEASIBILITY OF TROPOSPHERIC AND TOTAL OZONE DETERMINATION USING A FABRY-PEROT INTERFEROMETER AS A SATELLITE-BASED NADIR-VIEWING ATMOSPHERIC SENSOR

by

Allen Maurice Larar

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Atmospheric and Space Sciences)
in The University of Michigan
1993

Doctoral Committee:

Professor S. Roland Drayson, Chairperson
Associate Professor Mary L. Brake
Professor Paul B. Hays
Professor William R. Kuhn
Jack Fishman, Branch Head, NASA Langley Research Center
James M. Russell, III, Branch Head, NASA Langley Research
Center

N94-17214

Unclass

G3/46 0193061

(NASA-TM-109342) THE FEASIBILITY
OF TROPOSPHERIC AND TOTAL OZONE
DETERMINATION USING A FABRY-PEROT
INTERFEROMETER AS A SATELLITE-BASED
NADIR-VIEWING ATMOSPHERIC SENSOR
Ph.D. Thesis (Michigan Univ.)
194 p

ABSTRACT

THE FEASIBILITY OF TROPOSPHERIC AND TOTAL OZONE DETERMINATION USING A FABRY-PEROT INTERFEROMETER AS A SATELLITE-BASED NADIR-VIEWING ATMOSPHERIC SENSOR

by
Allen Maurice Larar

Chairperson: S. Roland Drayson

Monitoring of the global distribution of tropospheric ozone (O_3) is desirable for enhanced scientific understanding as well as to potentially lessen the ill-health impacts associated with exposure to elevated concentrations in the lower atmosphere. Such a capability can be achieved using a satellite-based device making high spectral resolution measurements with high signal-to-noise ratios; this would enable observation in the pressure-broadened wings of strong O_3 lines while minimizing the impact of undesirable signal contributions associated with, for example, the terrestrial surface, interfering species, and clouds. The Fabry-Perot interferometer (FPI) provides high spectral resolution and high throughput capabilities that are essential for this measurement task. Through proper selection of channel spectral regions, the FPI optimized for tropospheric O_3 measurements can simultaneously observe a stratospheric component and thus the total O_3 column abundance. Decreasing stratospheric O_3 concentrations may lead to an increase in biologically harmful solar ultraviolet radiation reaching the earth's surface, which is detrimental to health.

In this research, a conceptual instrument design to achieve the desired measurement has been formulated. This involves a double-etalon fixed-gap series configuration FPI along with an ultra-narrow bandpass filter to achieve single-order operation with an overall spectral resolution of approximately $.068\text{ cm}^{-1}$. A spectral region of about 1 cm^{-1} wide centered at 1054.73 cm^{-1} within the strong $9.6\text{ }\mu\text{m}$ ozone infrared band is sampled with 24 spectral channels. Other design characteristics include operation from a nadir-viewing satellite configuration utilizing a 9 inch (diameter) telescope and achieving horizontal spatial resolution with a 50 km nadir footprint.

A retrieval technique has been implemented and is demonstrated for a tropical atmosphere possessing enhanced tropospheric ozone amounts. An error analysis assessing the impact on retrieved O_3 amounts of the most significant uncertainties associated with this particular measurement has been performed for several different types of atmospheres. Results show the proposed instrumentation to enable a good measurement of absolute ozone amounts and an even better determination of relative changes, with a range of accuracy to within 7.55 to 20.6% for integrated tropospheric amounts (and 1.99 to 4.02% for total O_3 column abundance) and a corresponding range in precision to within 7.73 to 10.4% (and 3.30 to 3.95% for total O_3 column abundance), for the atmospheric conditions considered.

This thesis is dedicated to my parents for setting the foundation for me to make my goals such as this degree attainable.

ACKNOWLEDGEMENTS

Obtaining a Ph.D. degree is a long process requiring assistance from many people along the way. Acknowledgements provide an opportunity to let those individuals (even those inadvertently remaining unnamed) know that their help has been greatly appreciated. First, I would like to thank Professor S. Roland Drayson for all the advice and guidance that he as graciously provided as my advisor, a role for which I feel very fortunate to have had him serve. He has provided invaluable assistance in the development of my research and always managed to find time for our interactions, for which I am very grateful. I thank Professor Paul B. Hays for serving on my dissertation committee and for providing his Fabry-Perot interferometry expertise since the early phases of this research. I would also like to thank Professors Mary L. Brake and William R. Kuhn for serving on my committee and for their valuable comments and discussions.

I would like to acknowledge the NASA-Langley Research Center (NASA-LaRC) and the NASA Graduate Study Program for continuous support throughout my doctoral program at the University of Michigan. I would especially like to thank Dr. Jack Fishman, Dr. James D. Lawrence, Jr., Dr. James M. Russell, III, and Dr. Robert K. Seals, Jr., all of the Atmospheric Sciences Division at NASA-LaRC; Drs. Lawrence, Russell, and Seals for making possible this unique opportunity for me to pursue this advanced degree, and I thank Drs. Fishman and Russell for serving as special members on my doctoral committee.

I would like to recognize Dr. Jinxue Wang and Dr. Wilbert R. Skinner, of the Space Physics Research Laboratory at the University of Michigan; Dr. Wang for his many helpful discussions regarding various aspects of Fabry-Perot interferometry, and Dr. Skinner for sharing with me some of his knowledge pertaining to multiple-etalon configurations. Additionally, I would like to acknowledge Dr. Martin G. Mlynczak, of the Atmospheric Sciences Division at NASA-LaRC, for providing great advice regarding Professor Drayson and the University of Michigan, as well as for all the encouragement during my graduate studies.

Finally, without naming each one individually, I would like to thank all of my family members for their encouragement and understanding during this “busy” time period of graduate study. Most importantly, I owe a special thanks to my incredible wife Barbara for all her love, patience, and support throughout my degree program and always; and for making our memories of Ann Arbor that much more special with it being the birthplace of our wonderful son Jordan, who was my favorite study partner as I prepared for my preliminary examination to achieve doctoral candidacy.

TABLE OF CONTENTS

DEDICATION	ii
ACKNOWLEDGEMENTS	iii
LIST OF FIGURES	vii
LIST OF TABLES	xi
LIST OF APPENDICES	xiii
CHAPTER	
I. INTRODUCTION	1
II. SATELLITE-BASED PASSIVE DETECTION OF OZONE	9
2.1 Introduction	9
2.2 Atmospheric Ozone Vertical Distribution Characteristics	9
2.3 Selection of Spectral Band and Viewing Geometry	13
2.4 The Radiative Transfer Equation	19
2.5 Instrumental Function	24
III. APPLICATION OF FABRY-PEROT INTERFEROMETRY	26
3.1 Introduction	26
3.2 Spectral Isolation Using Multiple Etalons in Series	27
3.3 Methods of Tuning and the CLIO System	38
IV. DOUBLE-ETALON INSTRUMENT DESIGN TO ACHIEVE SINGLE-ORDER HIGH-RESOLUTION MEASUREMENT OF ATMOSPHERIC OZONE	46
4.1 Introduction	46
4.2 Water Vapor Continuum Absorption in the 10 μ m Spectral Region	47

4.3	Atmospheric and Total Upwelling Radiance Calculations . . .	51
4.4	Selection of Double-Etalon Instrument Characteristics	61
4.5	Band Placement and Choice of Spectral Channels	69
4.6	Further Design Aspects and Signal-to-Noise Ratio Analysis .	85
V. RETRIEVAL SIMULATION AND ERROR ANALYSIS . . .		96
5.1	Introduction	96
5.2	Ozone Retrieval Technique	96
5.3	Method of Error Analysis	100
5.4	Application to Atmospheric Cases	103
5.5	Retrieval of O_3 for a Perturbed Tropical Atmosphere	109
5.6	Results of Error Analysis	113
VI. SUMMARY AND CONCLUSIONS		132
APPENDICES		138
BIBLIOGRAPHY		171

LIST OF FIGURES

Figure

2.1	Atmospheric layers and temperature profile from 0 to 100 km, based upon the 1976 U.S. Standard Atmosphere.	10
2.2	The standard ozone profile from 0 to 100 km, of the 1976 U.S. Standard Atmosphere.	11
2.3	The normal vibrational and rotational modes of the nonlinear triatomic ozone molecule.	13
2.4	Schematic illustrating configuration for nadir-viewing satellite geometry.	15
2.5	The ozone 9.6 μm absorption band line strength distribution.	17
2.6	Approximate altitude dependence from 0 to 60 km of Lorentz, Doppler, and Voigt half-widths corresponding to ozone lines within the 9.6 μm band for the 1976 U.S. Standard atmosphere.	22
3.1	Transmission functions corresponding to two different double-etalon configuration scenarios having different plate spacer ratios.	29
3.2	A simple double-etalon series configuration using a slightly attenuating medium to reduce the undesirable effect of inter-etalon reflections.	32
3.3	Transmission character of a double-etalon system incorporating inter-etalon reflections while having an attenuating medium placed in between the etalons.	33
3.4	Schematic representation of the construction of a double-etalon FPI instrumental transfer function including broadening effects and incoherent reflections between etalons.	36
3.5	Qualitative diagram summarizing the formation of a double-etalon instrumental transfer function from the Airy (\mathcal{A}), defect (D), aperture (L), and bandpass filter (T_F) functions.	37

3.6	Conceptual diagram of the CLIO system optical configuration used to convert a circular Fabry-Perot fringe pattern into a linear pattern along the axis of a 45 degree half-angle internally reflecting transforming cone.	42
3.7	View from fringe plane of the CLIO system without the circle to line converter in place.	43
4.1	Representation of an inhomogeneous atmosphere by a sequence of homogeneous slabs.	52
4.2	Spectral lines located within the 9.6 μm ozone band according to the 1992 AFGL compilation.	56
4.3	Atmospheric transmission and upwelling radiance computed at .01 cm^{-1} resolution for the 1976 Standard Atmosphere.	57
4.4	Effect of atmospheric ozone perturbations on upwelling radiance within the 9.6 μm band region.	59
4.5	Effect of atmospheric ozone perturbations on upwelling radiance in a narrow spectral region of strong ozone lines.	62
4.6	White light transmitted through a double-etalon Fabry-Perot system.	65
4.7	Illustration of ultra-narrow bandpass filter function, broadened double-etalon transmission function, and overall instrumental transfer function of the central wavenumber for an FPI with the chosen specifications.	68
4.8	Spectral lines located within a portion of the 9.6 μm ozone band according to the 1992 AFGL compilation.	71
4.9	O ₃ energy contribution functions and how they are affected from including interfering species for the 1048-1058 cm^{-1} region.	72
4.10	Signal sensitivity to a 10 percent increase in tropospheric ozone.	73
4.11	Signal sensitivity to a 1 percent increase in stratospheric ozone.	74
4.12	Upwelling radiance reaching and received by the instrument within the selected band region.	76

4.13	O ₃ energy contribution functions and how they are affected from including interfering species for the instrument band spectral region.	77
4.14	Effect of atmospheric ozone perturbations on the radiance and signal within the instrument band spectral region.	78
4.15	Signal sensitivities to a 10 percent increase in tropospheric ozone and to a 1 percent increase in stratospheric ozone are shown with the selected instrument channels.	79
4.16	Energy Contribution Functions (ECFs) for the selected channels normalized to their peak values for an ozone only atmosphere.	80
4.17	Effect of including interfering species on ECFs from selected channels computed for an ozone only atmosphere.	82
4.18	Effect on O ₃ ECFs within the measurement band from including interfering species.	83
4.19	Schematic of proposed double-etalon instrument design configuration illustrating relative positioning of key optical elements.	86
4.20	Results from a signal-to-noise ratio (SNR) analysis using the U.S. Standard (1976) atmosphere.	92
5.1	Atmospheric models from the 1976 Standard Atmosphere database.	103
5.2	Ozone standard deviation profiles used in generating the a priori covariance matrices corresponding to atmospheric models within the 1976 Standard Atmosphere database.	105
5.3	Ozone and temperature profiles from the 1976 Standard Atmosphere (tropical) along with a sounding from Brazzaville, Congo, on August 5, 1992.	109
5.4	Retrieval simulation for a tropical atmosphere having enhanced tropospheric ozone.	112
5.5	Retrieved ozone sensitivity to perturbations of interfering species at specific altitudes for a tropical atmosphere.	116
5.6	Retrieved ozone sensitivity to temperature profile perturbations and systematic offsets in some interfering species for a tropical atmosphere.	117

5.7	Retrieved ozone sensitivity to perturbations in various forward model parameters for a tropical atmosphere.	120
5.8	Retrieved ozone sensitivity to perturbations in spectral line parameters and radiometric signal calibration errors for a tropical atmosphere.	121
5.9	Effects of component and total random errors on retrieved ozone for a tropical atmosphere.	126
A.1	Schematic diagram of monochromatic ray propagation through a simple Fabry-Perot interferometer system.	141
A.2	Ideal FPI throughput versus reflectivity for several values of plate absorption.	142
A.3	Transmission through ideal etalon for different plate reflectivities. .	143
A.4	Transmission through an ideal etalon with the corresponding FWHM (δ_F) and FSR (ν_{fsr}) depicted.	146
A.5	Reflectivity finesse versus reflectivity for an ideal etalon.	148
A.6	Common etalon plate imperfections and analytical functions frequently used to represent their corresponding defect distribution functions.	151
A.7	Etalon transmission characteristics with and without plate defects being present.	155

LIST OF TABLES

Table

4.1	Stratification used to represent the 1976 U.S. Standard atmosphere.	53
4.2	Simulation demonstrating effect of varying plate reflectivity while maintaining fixed spectral resolution on peak double etalon transmission $T_{D_{max}}$ and minimum white light W_{min} through the system.	66
4.3	Double-etalon instrument design specifications.	67
4.4	Characteristics of the O_3 ECFs for the instrument channels.	81
4.5	Orbital characteristics and dwell time determination assuming synchronization with the Nimbus-7/TOMS orbit.	91
4.6	Summary of detector element positioning and achievable SNR (assuming $D^* = 5.5 \times 10^{10} cm Hz^{\frac{1}{2}}/W$) for the chosen channels.	93
5.1	Summary of signal-to-noise ratio range estimates corresponding to the instrumentation of this research applied to different atmospheric types.	107
5.2	The effect of certain systematic perturbations on retrieved integrated ozone amounts for a tropical atmosphere.	119
5.3	The effect of systematic perturbations of the radiometric signal calibration on retrieved integrated ozone amounts for a tropical atmosphere.	122
5.4	Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a tropical atmosphere.	127
5.5	Summary of total (rss) errors in retrieved integrated O_3 amounts for all atmospheric cases considered.	129

A.1	Analytical expressions used to represents defect distribution functions for some common etalon plate defects.	152
A.2	Plate defect characteristics.	153
B.1	The effect of certain systematic perturbations on retrieved integrated ozone amounts for a mid-latitude summer atmosphere.	160
B.2	The effect of certain systematic perturbations on retrieved integrated ozone amounts for a mid-latitude winter atmosphere.	161
B.3	The effect of certain systematic perturbations on retrieved integrated ozone amounts for a subarctic summer atmosphere.	162
B.4	The effect of certain systematic perturbations on retrieved integrated ozone amounts for a subarctic winter atmosphere.	163
B.5	The effect of certain systematic perturbations on retrieved integrated ozone amounts for the U.S. standard (1976) atmosphere.	164
C.1	Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a mid-latitude summer atmosphere.	166
C.2	Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a mid-latitude winter atmosphere.	167
C.3	Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a subarctic summer atmosphere.	168
C.4	Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a subarctic winter atmosphere.	169
C.5	Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to the U.S. standard (1976) atmosphere.	170

LIST OF APPENDICES

Appendix

A.	THEORY OF FABRY-PEROT INTERFEROMETRY	139
	A.1 Introduction	139
	A.2 Characteristics of an Ideal FPI	140
	A.3 The Real Etalon	148
B.	ERROR ANALYSIS: SENSITIVITIES TO SYSTEMATIC PERTUR- BATIONS	159
C.	ERROR ANALYSIS: OVERALL ERROR BUDGETS	165

CHAPTER I

INTRODUCTION

In recent years, there has been an ever increasing need for improvements in the capabilities to remotely sound ozone (O_3) in the earth's atmosphere. This deepening interest has evolved from advances in the understanding of the significance of atmospheric ozone, as well as from recognition that anthropogenic activities may potentially affect the ozone spatial distribution and total column abundance.

Approximately 90 percent of ozone present in the earth's atmosphere is located in the stratosphere with maximum mixing ratios of around 10 parts per million by volume (ppmv) near 35 km. Most of the remaining ozone is situated in the troposphere where clean, background concentrations can range from 20 to 80 parts per billion by volume (ppbv) due to natural variability; however, tropospheric ozone concentrations can get as high as 500 ppbv in polluted urban areas (Seinfeld, 1986). Despite its relatively small abundance, atmospheric ozone plays a vital role in shielding life on earth from biologically harmful solar ultraviolet (UV) radiation; hence the term "good ozone" which is sometimes used to refer to stratospheric ozone. Conversely, tropospheric ozone is often called "bad ozone" since elevated surface concentrations can cause detrimental health effects, as well as damage to crops and other vegetation.

Many chemical species must be considered in order to understand the current-day

atmospheric ozone distribution and how it might change in the future. The man-made and natural chemicals of interest include the nitrogen oxides (NO_x) from urban surface emissions and high-altitude aircraft, nitrous oxide (N_2O) from agricultural and combustion practices, chlorofluorocarbons (CFC's) used primarily as aerosol propellants and refrigerants, bromine compounds, carbon monoxide (CO) and carbon dioxide (CO_2) from combustion processes, methane (CH_4) and other hydrocarbons which come from a variety of natural and anthropogenic sources. It has become well known that atmospheric concentrations of these species as well as other anthropogenically emitted gases have been increasing over recent years. This is of concern due to a potential climatic impact since these species in addition to O_3 and water vapor (H_2O) are significant in the energy budget and resulting thermal character of the atmosphere, and due to the potential of increased penetration of biologically harmful UV fluxes to the earth's surface resulting from reduced stratospheric ozone. Stratospheric ozone depletion occurs primarily through catalytic cycles involving hydrogen (HO_x), nitrogen (NO_x), chlorine (ClO_x), and to a lesser extent bromine (BrO_x) species.

Pollution produced by urban areas in the Northern Hemisphere is causing increased tropospheric ozone on a regional and possibly even global scale. Levels of tropospheric O_3 have been increasing and will continue to increase as concentrations of the precursor gases (oxides of nitrogen, methane and other hydrocarbons) necessary for the photochemical formation of tropospheric O_3 continue to rise. There is evidence suggesting that average surface O_3 concentrations have more than doubled in the past century at middle to high Northern Hemisphere latitudes (Volz and Kley, 1988). The finding of enhanced tropospheric ozone over the south tropical Atlantic Ocean (Fishman et al., 1990), in a region remote from industrialized areas, shows

that the problem is not unique to urban areas; this situation appears to result from transport of precursors emitted from wide spread vegetation burning primarily in Africa, and the subsequent photochemical production of ozone.

Information on the polar ozone distribution has become increasingly desirable during the last several years. The unique meteorological conditions present in polar regions enable heterogeneous chemistry, involving ice crystal surfaces within Polar Stratospheric Clouds (PSC's), to enhance the catalytic destruction of stratospheric ozone (primarily by man-made chlorine species) leading to massive depletions in the polar springtime; this was first realized with the initial reporting of the "ozone hole" over Antarctica in 1985 (Farman et al., 1985; Stolarski et al., 1986). In addition, global total ozone is also decreasing, with rates having a geographical location, altitude, and seasonal dependence (Watson et al., 1988). In fact, there is strong observational evidence from many different measurement techniques showing that, over the past decade, annual average ozone has decreased by approximately 10 percent in the mid- and high-latitude lower stratosphere (WMO, 1991). The interest regarding ozone depletion over sparsely-populated polar regions is not simply due to scientific curiosity but, more importantly, the concern of whether similar depletions can occur in the more densely populated lower latitudes; this might be possible from dynamical transport of ozone depleted air from the polar regions (through the vortex or after the vortex breaks down during the polar springtime), or possibly from an analogous heterogeneous chemical mechanism involving stratospheric aerosol surfaces.

Knowledge of the global scale distribution of atmospheric ozone and its temporal variability can only be achieved through implementation of a satellite-based remote sensing technique. Many satellite-based passive instrumentation methods (having the radiation source external and thus not provided by the instrument) have been

previously implemented primarily to infer stratospheric ozone amounts and, to a lesser extent, total column abundances of atmospheric ozone (Grant, 1989). These instruments have included absorption techniques involving direct and indirect solar radiation as well as emission measurements, and have operated in both limb and nadir viewing satellite geometries. For example, the Stratospheric Aerosol and Gas Experiment (SAGE) I&II instruments (McCormick et al., 1979; McCormick, 1987) are satellite-based spectrometers that measure the absorption of solar UV radiation by ozone, and operate in a limb-viewing solar occultation geometry by acquiring data only during spacecraft sunrise and sunset events. The Halogen Occultation Experiment (HALOE) instrument (Russell et al., 1993; Baker et al., 1986) also infers stratospheric ozone vertical profiles by observing absorption of solar radiation from a solar occultation viewing geometry, however, this is accomplished using a broadband filter radiometer operating in the IR portion of the spectrum. There has been a series of instruments employed to deduce ozone amounts through measurement of backscattered solar UV radiation observing the terrestrial atmosphere from a nadir-viewing geometry. Most recently, the Solar Backscatter Ultraviolet (SBUV 2) instrument (Frederick et al., 1986) has been used to infer total ozone as well as for ozone profiling. The Total Ozone Mapping Spectrometer (TOMS) instrument (Heath et al., 1975) also measures backscattered solar UV radiation similar to an SBUV instrument, but it allows for off-nadir viewing by scanning cross-track and it provides integrated ozone amounts rather than profiles. Alternatively, atmospheric ozone amounts can be deduced from observation of molecular emission. The Infrared Interferometer Spectrometer (IRIS) was a Michelson interferometer instrument employed in a nadir-viewing configuration to measure IR emission by ozone in the 9.6 μm band spectral region (Hanel et al., 1971). Operating in a limb-viewing geometry,

the Limb Infrared Monitor of the Stratosphere (LIMS) instrument (Gille and Russell, 1984) employed a broadband filter radiometer to also measure emission by ozone in the IR ($9.6\ \mu\text{m}$ band) spectral region.

Many different types of instrumentation have been designed for the measurement of stratospheric and total ozone, however, development of a satellite-based tropospheric ozone remote sensor has remained a difficult task to accomplish due to the challenges associated with such a measurement. Several factors contribute to this difficulty. For instance, clouds usually restrict a limb-viewing technique from sensing below the upper troposphere, as the long horizontal path is very likely to contain cloud interferences in the instrument's field-of-view (FOV). This effect can be minimized by the small horizontal footprint (and thus better horizontal resolution) associated with a nadir-viewing configuration, however, then one usually has to contend with a large surface contribution to the measured signal. In addition, when viewing in the nadir, the necessary vertical resolution is usually achieved by raising the spectral resolution; although, this directly negatively impacts the maximum attainable signal-to-noise of the measurement. Another consideration in sensing the lower atmosphere is that the unfavorable contribution of interferant species can be worse, as the increased atmospheric pressure causes their overlap effects to be spectrally broader or possibly continuum in nature.

Although a satellite-based instrument capable of directly measuring tropospheric ozone has never been developed, data from existing instruments have been used to deduce characteristics associated with the global tropospheric ozone field. The TOMS instrument is used to measure total column ozone, however, its signal can be sensitive to low altitude ozone variations (Fishman and Browell, 1988); therefore, its measurements can be used to infer variations in tropospheric ozone (Fishman et

al., 1987). In addition, TOMS data have been used in conjunction with vertically integrated SAGE profile data to deduce tropospheric “residual” amounts (Fishman and Larsen, 1987). This method requires the spatial mapping of coincident data and associated measurement errors from two different instruments.

An instrument capable of observing tropospheric ozone must be of high spectral resolution in order to isolate the pressure-broadened wings of strong O_3 lines and have sufficient signal-to-noise to enable removal of effects introduced by the surface and interferant species. A Fabry-Perot Interferometer (FPI) has a high optical throughput capability while operating at high spectral resolution and has the possibility of being very advantageous when compared with conventional spectrometers (Hernandez, 1986; Jacquinet, 1954; Vaughan, 1989). In addition, a multiplex advantage is achievable as all necessary frequencies can be measured simultaneously using a multichannel configuration. These characteristics make the FPI a suitable candidate for attempting this difficult tropospheric remote sensing measurement. Instruments utilizing Fabry-Perot interferometry have been used to monitor the earth’s atmosphere from the surface for some time and have also been involved in previous satellite-based applications (Vaughan, 1989). A spherical Fabry-Perot interferometer was used to measure the 630 nm airglow oxygen emission line aboard the Orbiting Geophysical Observatory (OGO-6) during 1969 and 1970 (Blamont and Luton, 1972); upper atmospheric temperatures of neutral components in the thermosphere were then deduced from these measurements. A plane Fabry-Perot interferometer was launched on the Dynamics Explorer (DE) satellite in 1981 to measure the temperature, meridional wind and density of several species of excited atoms and ions in the thermosphere (Hays et al., 1981). The High Resolution Doppler Imager (HRDI) is a triple etalon Fabry-Perot interferometer (Skinner et al., 1987), aboard the Upper

Atmosphere Research Satellite (UARS) launched in 1991, designed to infer upper atmospheric wind vectors by observing the Doppler shift of atmospheric absorption and emission lines in the O_2 atmospheric bands (Hays, 1982). Also on the UARS platform is the Cryogenic Limb Array Etalon Spectrometer (CLAES) instrument (Roche and Kumer, 1989) which is a spectrometer utilizing solid Fabry-Perot etalons; the primary objective of CLAES is to retrieve stratospheric profiles of temperature and trace species related to ozone layer photochemistry (including ozone itself) from measurement of limb IR spectral emissions. Both the DE-FPI and the HRDI instrument were designed to perform a multichannel measurement by recording fringes on image plane detectors consisting of concentric ring anode elements positioned in their fringe planes. These detectors enable simultaneous monitoring of all spectral channels and utilize the entire fringe to achieve maximum signal. The satellite-based Fabry-Perot instruments thus far which have utilized the entire fringe pattern have been limited to the visible portion of the spectrum. Similar applications at longer wavelengths have not been exploited due in part to lack of detector technology in the IR spectral region. The IR-FPI detection difficulty will be overcome in this research through implementation of the new Circle to Line Interferometer Optical (CLIO) system (Hays, 1990; U.S. patent #4893003, 1990), developed by researchers at the Space Physics Research Laboratory (SPRL) of the University of Michigan. This system converts the circular interferometric fringes into a linear pattern which then can be detected by conventional linear array detectors; and thus opening the door for a host of possible IR-FPI applications.

The intent of this research is to assess the feasibility of determining tropospheric and total ozone using a single-order double-etalon Fabry-Perot interferometer operating from a satellite-based nadir-viewing geometry. The basic characteristics of

ozone will be presented in chapter II, along with the radiative transfer equation for upwelling radiation reaching a satellite instrument configured in the nadir-viewing geometry. Chapter III and supporting appendix A contain the necessary theory of Fabry-Perot interferometry, starting from a simple, ideal instrument and progressing to real instrument characteristics in a multi-etalon configuration; chapter III concludes with a brief discussion of tuning techniques while having a concentration on the new CLIO system. A development of the proposed double-etalon FPI instrument design to achieve a single-order high-resolution measurement of atmospheric ozone is presented in chapter IV. In chapter V, a retrieval technique and error analysis procedure are formulated and then applied to realistic atmospheric test cases to simulate instrument performance capabilities. A summary of this research as well as suggestions for future related efforts are given in the concluding chapter VI.

CHAPTER II

SATELLITE-BASED PASSIVE DETECTION OF OZONE

2.1 Introduction

The only practical means of obtaining knowledge pertaining to the global distribution of trace atmospheric species and their associated temporal variability is through implementation of a satellite-based remote sounding measurement technique. Ozone is the atmospheric constituent of interest for this research endeavor and, accordingly, the emphasis of this chapter will be directed toward space-based detection of ozone. The chapter will discuss ozone characteristics, including atmospheric spatial distributions and spectral bands, as well as present the radiative transfer equation for the desired nadir-viewing satellite geometry. In concluding, this chapter will formulate expressions for measured signal, weighting, and energy contribution functions while incorporating the instrumental function to account for the measurement process.

2.2 Atmospheric Ozone Vertical Distribution Characteristics

The atmosphere is conventionally divided up into layers based upon the vertical structure of its temperature distribution. The approximate positioning of these layers, the troposphere, stratosphere, mesosphere, and thermosphere is illustrated

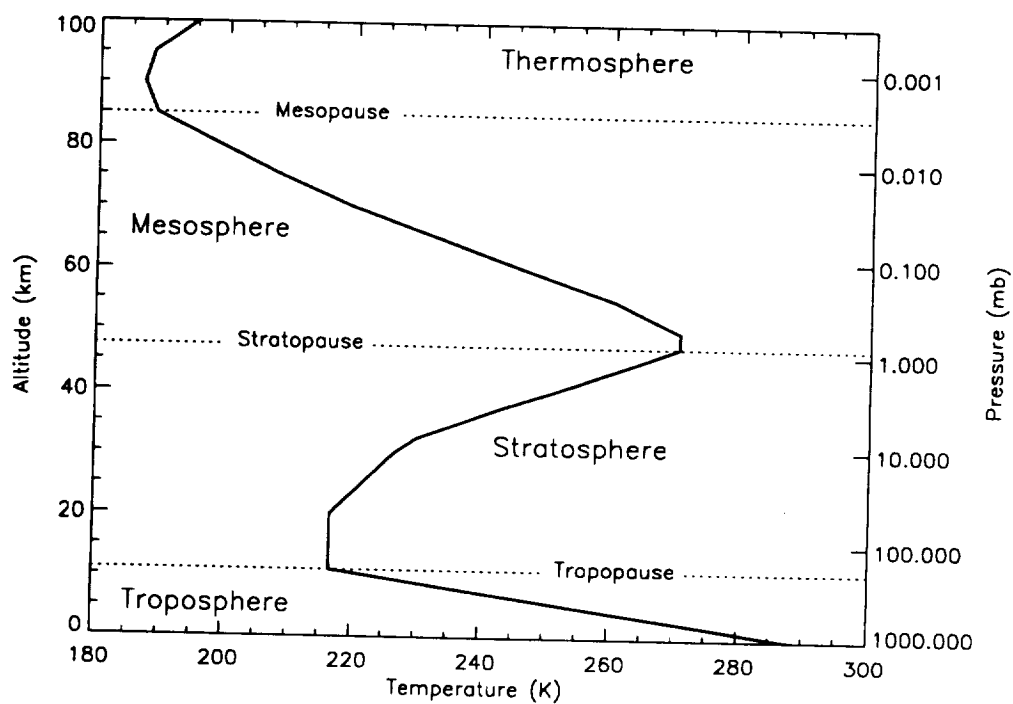


Figure 2.1: Atmospheric layers and temperature profile from 0 to 100 km, based upon the 1976 U.S. Standard Atmosphere.

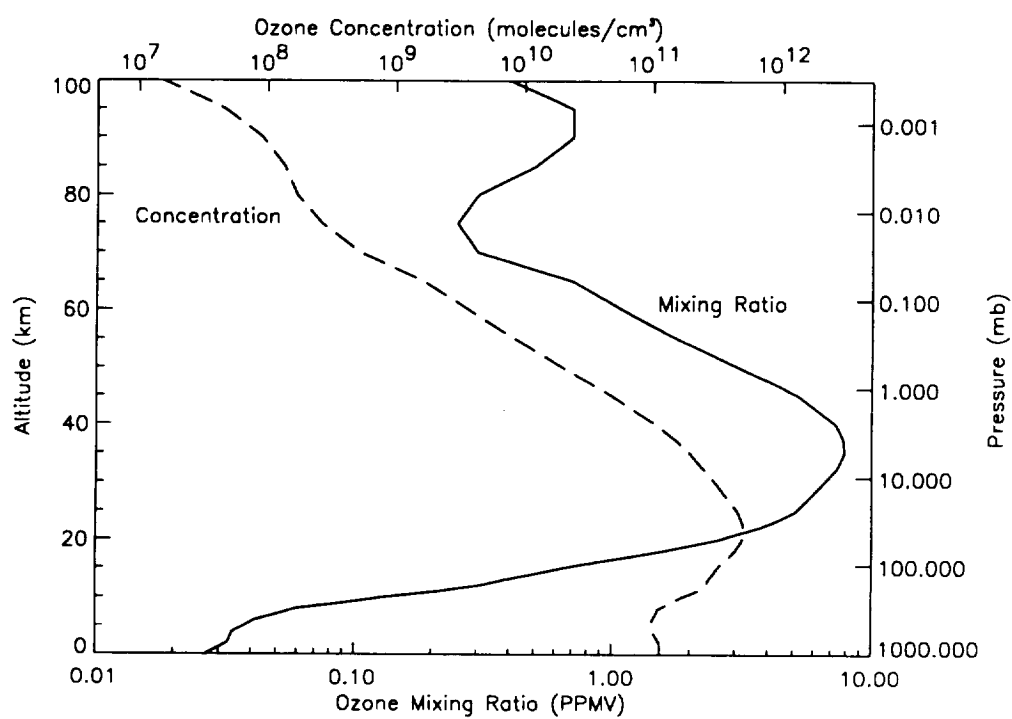
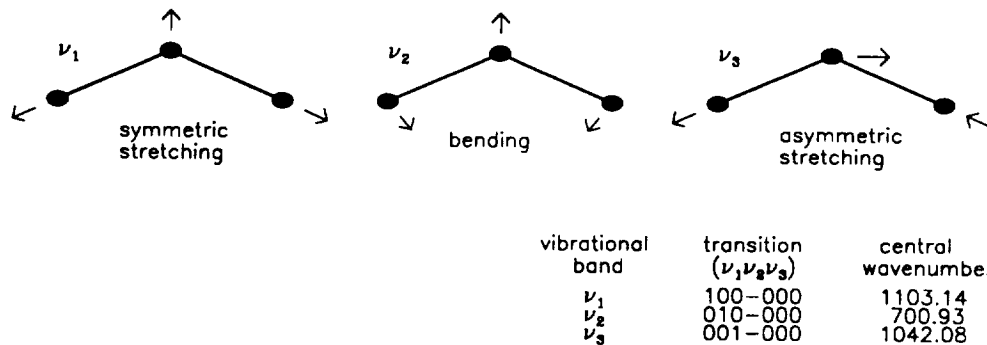


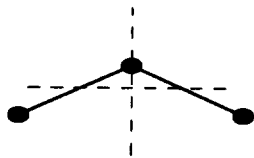
Figure 2.2: The standard ozone profile from 0 to 100 km, of the 1976 U.S. Standard Atmosphere.

in Figure 2.1. The average temperature distribution shown is based upon the 1976 Standard Atmosphere data set (Anderson et al., 1986), and has its vertical character resulting from radiative processes involving both solar and terrestrial radiation. The rise in temperature in the stratosphere to a peak at the stratopause is due to the absorption of solar ultraviolet radiation by ozone. The vertical profile of average ozone for the standard atmosphere (U.S. Standard (1976), Anderson et al., 1986) is plotted in Figure 2.2 in units of molecular concentration and mass mixing ratio. As this Figure illustrates, most of the mass of atmospheric ozone is contained in the lower stratosphere with a maximum concentration found at an altitude of around 20 km; whereas a maximum in ozone mixing ratio occurs much higher, near 35 km. The plots shown correspond to average vertical structure; the altitude of maximum ozone for an individual profile varies with season, latitude, slightly with longitude, and meteorological conditions. Approximately 90 percent of the ozone mass resides in the stratosphere, around 10 percent in the troposphere, and the residual trace amount can be found above the stratopause. Ozone extends over a large vertical range. However, as even the magnitude of the peak ozone mixing ratio demonstrates, ozone is indeed a trace constituent; nevertheless, it is the most significant trace species in the middle atmosphere. Its extremely efficient absorption of UV radiation is realized when one considers if all the ozone in a column of air extending from the surface to the top of the atmosphere were reduced to standard temperature and pressure (0°C , 1013.25 mb), the column height would be only around 3mm! Total ozone is defined as being the total amount of ozone present in a vertical column and is commonly measured in Dobson units (DU). A Dobson unit is defined by $1\text{DU} = 10^{-3}\text{cm}$ (or $1\text{DU} = 2.69 \times 10^{16}$ molecules $\text{O}_3 \text{ cm}^{-2}$); therefore, the global mean abundance of ozone is approximately 300 DU. The main isotopic species of ozone $^{16}\text{O}_3$ accounts for

Normal Vibrational Modes:



Normal Rotational Modes:



3 rotational modes about axes of 3-d coordinate system centered at center of mass of molecule

Figure 2.3: The normal vibrational and rotational modes of the nonlinear triatomic ozone molecule.

over 99 percent of atmospheric ozone. Ozone can also contain an atom of heavy oxygen ^{18}O yielding the minor isotopic species $^{16}O^{16}O^{18}O$ and $^{16}O^{18}O^{16}O$ which make approximately .4% and .2%, respectively, of atmospheric ozone (Khrigian, 1975).

2.3 Selection of Spectral Band and Viewing Geometry

Molecules can have rotational (due to rotations of the nuclei), vibrational (due to nuclear vibrations), or electronic (due to motion of the electrons about the nuclei) energies, which individually and thus as a sum total can exhibit only discrete values; thus they are quantized. In addition a molecule can have translational energy, a continuous form of energy due to motion of the molecule's center of mass in space. A loss or gain in energy $E = h\omega$ occurs when a molecule emits or absorbs a photon of frequency $\omega(s^{-1})$, respectively. This forms the basis for the existence of line spec-

tra, since a molecule can absorb or emit only at discrete frequencies corresponding to a change from one quantized energy level to another. Although these molecular energies can only change by discrete amounts, numerous energy transitions are possible corresponding to different combinations of changes among the electronic, vibrational, and rotational energy levels. However, for a particular wavelength region certain transitions may not be possible. For example, electronic energy transitions typically correspond to large energy differences occurring at wavelengths less than $1\mu\text{m}$ and are therefore found in the visible and UV spectral regions. Whereas small energy differences are associated with pure rotational transitions and thus occur at longer wavelengths ($20\mu\text{m} < \lambda < 1000\mu\text{m}$) located in the far-IR portion of the spectrum. And the infrared region $\lambda < 20\mu\text{m}$ is where one would expect to find vibrational transitions. Of interest in explaining the infrared radiative transfer of atmospheric ozone are the so called “vibrational-rotational bands.” As its name implies, this type of band results from a simultaneous change in rotational energy that often occurs when a molecule undergoes a change in vibrational energy level. The occurrence of these transitions and thus the corresponding spectral bands can be better appreciated by considering molecular structure. Ozone is a nonlinear triatomic molecule with its atoms situated at vertices of an isosceles triangle (Khrghian, 1975). It possesses a relatively strong rotational structure, as well as three fundamental vibrational bands, ν_1 , ν_2 , and ν_3 at 9.066, 14.27, and 9.597 μm , respectively. The normal vibrational and rotational modes and resulting spectral bands for the ozone molecule are illustrated in Figure 2.3. The very strong ν_3 and fairly strong ν_1 fundamentals combine to form the important 9.6 μm band of ozone. The ν_2 fundamental at 14 μm is well-masked by the intense 15 μm band of CO_2 . Another strong ozone band near about 4.7 μm is located in a spectral region of low available atmo-

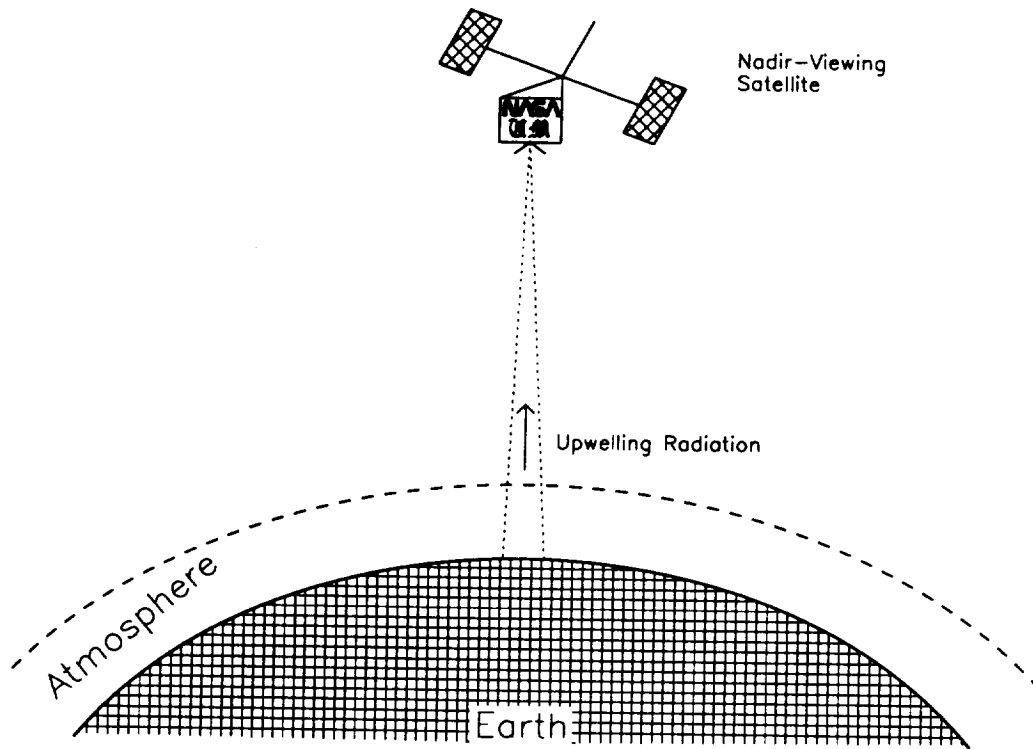


Figure 2.4: Schematic illustrating configuration for nadir-viewing satellite geometry.

spheric radiation, positioned in between the solar and terrestrial emission curves. It is also worth noting that ozone's absorption bands in the UV part of the spectrum are the result of electronic transitions of the ozone molecule, as expected, since electronic transitions usually occur at shorter wavelengths ($\lambda < 1 \mu\text{m}$) in the visible and ultraviolet portions of the spectrum.

The optical properties of ozone, namely its characteristics of absorption and emission of radiation, are responsible for shielding the biosphere from potentially harmful solar UV energy which results in formulating the stratospheric temperature profile shape; however, most importantly from a measurement viewpoint, these optical properties provide the basis for remote detection of its presence. Atmospheric sensing measurements by satellite instruments utilize two distinct kinds of viewing geometry: the limb view where radiation leaving the atmosphere on a nearly tangential

trajectory is observed, and the near-nadir view in which the instrument's field-of-view (FOV) observes radiation leaving the atmosphere in directions almost perpendicular to the surface. The limb view generally enables high vertical resolution (determined by the instrument's FOV), however, a coarse horizontal resolution (dictated by concentration distributions and radiative characteristics of the target species as well as atmospheric temperature) is usually obtained. In addition, sensing below the upper troposphere is usually not possible due to FOV obstructions by clouds. The smaller horizontal footprint of a nadir view (determined by the instrument's FOV) can minimize the limiting effect of clouds in viewing the troposphere. However, the vertical resolution in nadir viewing can be somewhat coarse as, like the limb viewing horizontal resolution, it is determined by (distribution, thermal, and spectral) properties of the target species. Also, knowledge of surface characteristics (such as temperature and emissivity) are usually required to permit removal of the lower boundary signal contribution. To enable sensing the troposphere with a relatively small horizontally footprint while minimizing the impact of clouds, a nadir-viewing geometry (as illustrated in Figure 2.4) will be assumed for this research application. When applying this remote sensing technique, one utilizes measurements of radiation emerging from the top of the atmosphere in several spectral intervals (or channels) ranging from high absorption to low absorption within the band of interest. Thermal energy observed in the high absorption region originates from high altitudes, as emission from lower levels gets absorbed by the atmosphere; conversely, observations in lower absorption spectral regions enable viewing lower in the atmosphere. However, relatively high stratospheric ozone concentrations further complicate the task of sensing tropospheric ozone. Nevertheless, with knowledge of the atmospheric temperature profile, measured spectral variation of outgoing radiation can be used

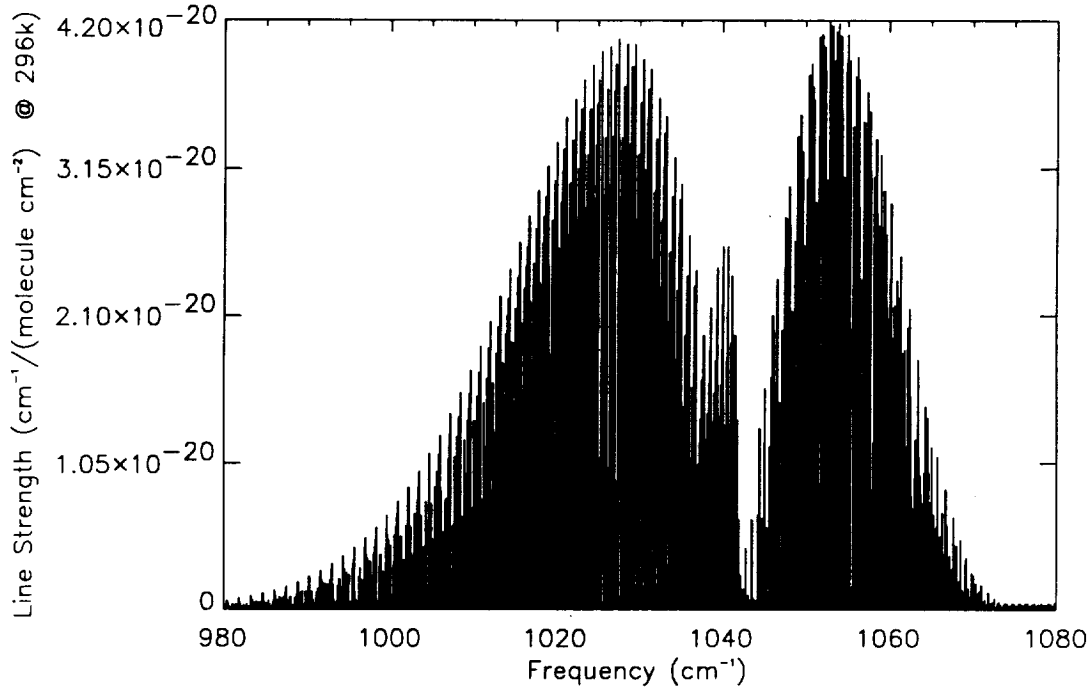


Figure 2.5: The ozone 9.6 μm absorption band line strength distribution.

to infer information on the absorbing gas distribution.

Ozone possesses absorption bands at various positions in the electromagnetic spectrum. While reflected solar UV radiation could be measured, continuous observations are possible during day (independent of solar zenith angle) or night by sensing in the infrared spectral region. As already discussed, ozone has several vibration-rotation bands in the infrared. In particular, some of its stronger bands have absorption maxima centered about 4.75, 9.57, and 14.2 μm . As mentioned by Houghton et al. (1984), several factors should be considered prior to selecting an absorption band for a remote sounding application. The effect of interferant species should be minimized by choosing a band that is not overlapped by bands of other atmospheric constituents; this would push in favor of the 9.6 μm band due to the presence of strong CO_2 bands that occur near wavelengths of 4.3 and 15 μm . It is desirable

for local thermodynamic equilibrium (LTE) to apply, so that the source function for thermal emission from the band will equal the Planck function, and thus simplify the radiative transfer calculations. LTE is said to exist for a particular transition when the ratio of number of molecules in the excited energy state to those in the ground state is given by a Boltzmann distribution at the local kinetic temperature; this will be the case if the probabilities of excitation and de-excitation by collisional processes are significantly larger than the probabilities of excitation and de-excitation by any radiative or chemical processes, and therefore decreases in likelihood at higher altitudes. From LTE considerations, it is best to view emission in a spectral region near terrestrial blackbody emission (8-12 μm), since terrestrial radiation is driving the molecules to excitation temperatures near their local atmospheric kinetic temperatures. Conveniently, ozone has a strong absorption/emission band in this spectral window region. It has been estimated (Mlynczak, 1989) that LTE is a valid assumption for the 9.6 μm ozone band for heights up to slightly lower than 60 km and 70 km for daytime and nighttime, respectively; since signal contributions at or above these altitudes are negligible for a nadir-viewing instrument sensing in the 9.6 μm spectral region, it would be safe to neglect non-LTE effects for such an application. It is also beneficial for the wavelength of the chosen band to be long enough such that scattered solar radiation is insignificant and thus can be neglected relative to thermal emission; this is the case for wavelengths greater than approximately 4 to 5 μm , which makes the $\lambda = 4.75 \mu\text{m}$ band a little less desirable. In addition, it is preferable to observe a strong signal from the target species of interest. This signal is formed by a complex coupling between vertical concentration profiles of the radiating species and the vertical temperature field, characteristics of which are determined by the particular atmospheric vertical region of interest and associated spatial and

temporal variability, as well as the strength of the absorption band. The distributions of species and temperature are not controllable, however, some freedom may exist in band selection. Actually, from a maximum signal standpoint it is desirous for one to choose a band having strong absorption lines located spectrally near the peak blackbody emission for the ambient temperature of the atmospheric region to be sensed; this roughly corresponds to 10 to 14 μm for the warm to cold temperature range, respectively, of the U.S. Standard atmosphere average temperature within the troposphere and stratosphere. In the IR portion of the spectrum, the strongest ozone bands are near 10 μm , where absorption results primarily from the ν_3 and ν_1 fundamentals of $^{16}\text{O}_3$ (Flaud et al, 1990). All of these factors provide a propensity toward basing this measurement on the $\lambda = 9.6 \mu\text{m}$ (wavenumber 1043 cm^{-1}) ozone band, which is a narrow but strong infrared absorption band consisting of a number of closely situated spectral lines. The spectral line distribution of this band (Rothman et al., 1987; Flaud et al., 1990) are illustrated in Figure 2.5.

2.4 The Radiative Transfer Equation

The radiative transfer equation (RTE), describing the transfer of radiant energy through the atmosphere, can be formulated uniquely for the particular satellite viewing geometry. For a nonscattering atmosphere in LTE, the RTE expressing clear column (cloud free) radiance reaching a nadir-viewing instrument at wavenumber ν and zenith angle θ may be written as

$$I(\nu, \theta) = \epsilon_s(\nu, \theta)B(\nu, T_s)\tau(\nu, \theta, p_s) + \int_{\ln p_s}^{-\infty} B(\nu, T(p))\frac{\partial \tau(\nu, \theta, p)}{\partial \ln p} d \ln p; \quad (2.1)$$

$\epsilon_s(\nu, \theta)$ represents the surface emissivity, $T(p)$ is the atmospheric vertical temperature profile as a function of pressure, $\tau(\nu, \theta, p)$ corresponds to the atmospheric spectral transmittance between pressure level p and the instrument, and the subscript s de-

notes a surface parameter. In addition, the radiation emitted by a blackbody at wavenumber ν and temperature T is described by the Planck function $B(\nu, T)$ and is given by

$$B(\nu, T) = \frac{a\nu^3}{\exp\left(\frac{b\nu}{T}\right) - 1}, \quad (2.2)$$

with the constants defined by $a = 2hc^2$ and $b = \frac{hc}{k}$; c is the speed of light, while h and k are Planck's and Boltzmann's constants, respectively. Equation 2.1 assumes reflected solar energy to be negligible, since the solar energy reaching the earth's atmosphere (solar Planck function with "dilution" factor) is insignificant relative to that of terrestrial origin for the spectral region of application. In addition, the contribution from surface reflected downward thermal flux is ignored for this study; this assumes the bottom boundary to be well represented by a blackbody, which is not a bad assumption since IR surface emissivities are usually close to unity. However, note that there are exceptions to the validity of this assumption; for example, over desert regions where the surface emissivity may be $\epsilon_s \ll 1$, for which cases inclusion of the surface reflected downward thermal flux may be necessary. The first term in the RTE (equation 2.1) represents the contribution from surface emission while the second term is due to atmospheric emission. The quantities $\frac{\partial \tau(\nu, \theta, p)}{\partial \ln p}$ and $B(\nu, T(p)) \frac{\partial \tau(\nu, \theta, p)}{\partial \ln p}$ are frequently referred to as weighting functions and (energy) contribution functions, respectively, as they indicate the vertical dependence of the information content contained within the upwelling radiance leaving the top of the atmosphere. The monochromatic transmittance $\tau(\nu, \theta, p)$ can be given by

$$\tau(\nu, \theta, p) = \exp\left[-\int \sum_i \kappa_{\nu_i} \rho_i ds\right] = \exp\left[-\int \sum_i \kappa_{\nu_i} du_i\right], \quad (2.3)$$

where κ_{ν_i} and ρ_i are the monochromatic absorption coefficient and density, respectively, for species i , ds is an increment in path length, and $du_i \equiv \rho_i ds$ is the com-

monly used increment of masspath. The summation over the index i accounts for the decrease in transmission due to multiple species in the path and arises from monochromatic transmittances being multiplicative; for example, the transmittance through a layer of N absorbing species ($\tau = \tau_1 \tau_2 \cdots \tau_N$) can be expressed as:

$$\tau = \exp \left[- \int \kappa_{\nu_1} du_1 \right] \exp \left[- \int \kappa_{\nu_2} du_2 \right] \cdots \exp \left[- \int \kappa_{\nu_N} du_N \right] = \exp \left[- \int \sum_i^N \kappa_{\nu_i} du_i \right]. \quad (2.4)$$

The transmittance through a series of consecutive homogenous paths can be obtained by multiplying the monochromatic transmittance for each path. Transmittances through inhomogeneous paths (such as through the earth's atmosphere) may be approximated by a series of thin homogeneous paths (i.e., Weinreb and Neuendorffer, 1973).

The physical structure of molecules determines their energy levels associated with possible transitions, and thus the central wavenumber about which the molecules absorb and emit radiation. Whereas, the applicable broadening mechanism dictates the spectral distribution of the particular absorption/emission feature about its central wavenumber. Correspondingly, the absorption coefficient of a spectral line can be expressed as (Kuhn and London, 1969):

$$\kappa_{\nu} = SP(\nu - \nu_0), \quad (2.5)$$

where S is the line strength (or integrated absorption coefficient) which is defined by $S = \int_0^{\infty} \kappa_{\nu} d\nu$, ν_0 is the central (or line center) wavenumber, and $P(\nu - \nu_0)$ is the normalized broadening profile function ($\int_0^{\infty} P(\nu - \nu_0) d\nu \equiv 1$). The shape of the profile function is usually determined by two broadening mechanisms for radiation in the earth's atmosphere. Molecular collisions during the absorption/emission process give rise to collisional or pressure broadening, and the corresponding profile function

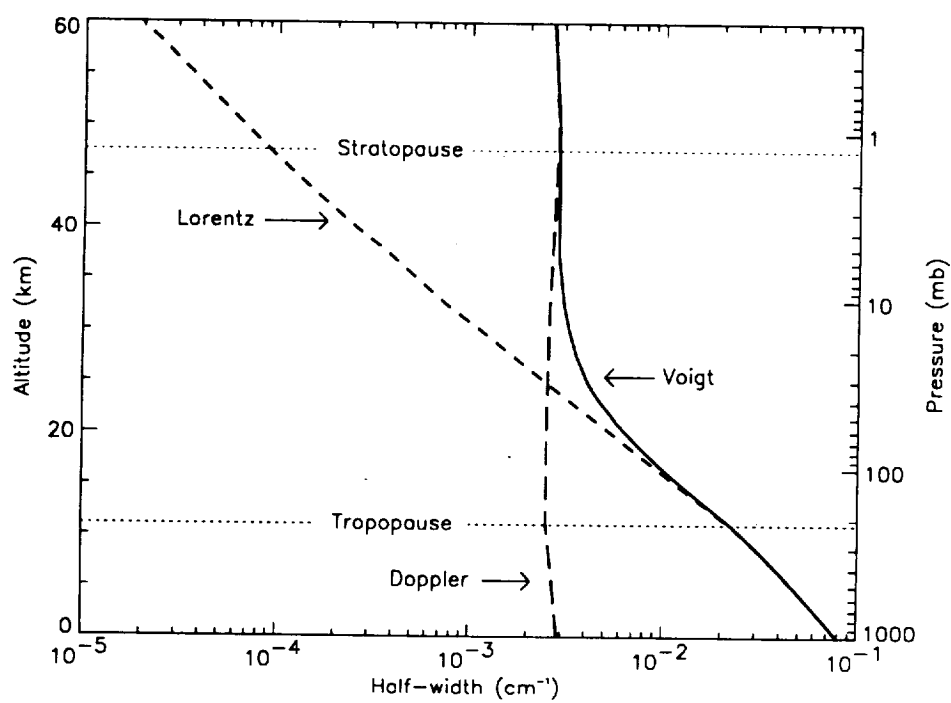


Figure 2.6: Approximate altitude dependence from 0 to 60 km of Lorentz, Doppler, and Voigt half-widths corresponding to ozone lines within the $9.6 \mu\text{m}$ band for the 1976 U.S. Standard atmosphere.

can be represented by the Lorentz line shape as

$$P_L(\nu - \nu_0) = \frac{1}{\pi} \frac{\alpha_L}{\alpha_L^2 + (\nu - \nu_0)^2}; \quad (2.6)$$

α_L is the Lorentz half-width, and represents half the width at half the maximum of the profile function. The Lorentz half-width is proportional to pressure and is approximately inversely proportional to the square root of temperature. Actually, $\alpha_L \propto \frac{p}{T^n}$ where n ranges from .5 to 1.0 (Rothman et al., 1987). The other significant mechanism for broadening in the terrestrial atmosphere is Doppler broadening which originates from a Doppler shift in the frequency of radiation associated with the absorption/emission feature due to thermal motion of the radiating molecules. The Doppler broadening profile function is of Gaussian form and can be given by

$$P_D(\nu - \nu_0) = \frac{1}{\alpha_D} \sqrt{\frac{\ln 2}{\pi}} \exp \left[-\ln 2 \left(\frac{\nu - \nu_0}{\alpha_D} \right)^2 \right], \quad (2.7)$$

where the Doppler half-width can be represented analytically as $\alpha_D = \sqrt{\frac{2kT \ln 2}{mc^2}} \nu_0$; in this expression, k is Boltzmann's constant, T is the temperature in Kelvin, m is the mass of the radiating molecule, and c is the speed of light. Unlike the Lorentzian half-width, the Doppler half-width does not have a pressure dependence and therefore its change with altitude is due to temperature alone. As discussed by Kuhn and London (1969), at low altitudes ≤ 20 km (high pressures) pressure broadening dominates and at high altitudes ≥ 50 km (low pressures) Doppler broadening is the most significant. However, for intermediate altitudes α_L and α_D can be comparable in size, at which both broadening mechanisms are significant and must be considered simultaneously. The resulting broadening profile function is formed by the convolution of the Lorentzian and Doppler line shapes yielding the so called Voigt profile which can be expressed as (see for example, Armstrong, 1967; Drayson, 1976):

$$P_V(\nu - \nu_0) = \frac{1}{\alpha_D} \sqrt{\frac{\ln 2}{\pi}} \frac{y}{\pi} \int_{-\infty}^{\infty} \frac{\exp(-t^2)}{y^2 + (x - t)^2} dt, \quad (2.8)$$

where the ratio of Lorentz to Doppler widths is defined by $y = \frac{\alpha_L}{\alpha_D} \sqrt{\ln 2}$ and $x = \frac{\nu - \nu_0}{\alpha_D} \sqrt{\ln 2}$ corresponds to a wavenumber scale in units of Doppler width. The Voigt profile is formed from the combination of two independent broadening formulas, and in the high pressure ($y \rightarrow \infty$) or low pressure ($y \rightarrow 0$) limits it approaches the Lorentz or Doppler profile, respectively. Figure 2.6 illustrates the approximate Lorentz, Doppler, and Voigt half-widths as a function of altitude for O_3 lines within the $9.6 \mu\text{m}$ band; $\alpha_{L,} = .08 \text{ cm}^{-1}$ and $\alpha_{D,} = 2.89 \times 10^{-3} \text{ cm}^{-1}$ (Kuhn and London, 1969) were used for surface ($T = 288 \text{ K}$ and $p = 1 \text{ atm}$) values of the Lorentz and Doppler half-widths, respectively, and $n = .76$ was used for the Lorentz half-width temperature dependence coefficient (Rothman et al., 1987). Note that regardless of the dominant broadening mechanism, the far-wing line shape is often represented simply by a Lorentzian line shape. The actual far-wing shape may depart from Lorentzian, but this will not be of concern for the current application since measurement channels will not extend too far from line centers in the wings of strong lines (i.e., for the tropospheric O_3 observations).

2.5 Instrumental Function

The RTE (equation 2.1) gives an expression to evaluate upwelling atmospheric and surface radiance $I(\nu, \theta)$, the radiation leaving the top of the atmosphere and reaching a satellite sensor, and contains the quantities $\frac{\partial \tau(\nu, \theta, p)}{\partial \ln p}$ and $B(\nu, T(p)) \frac{\partial \tau(\nu, \theta, p)}{\partial \ln p}$, the weighting and energy contribution functions, respectively, which indicate the vertical dependence of the contributions to this radiance. However, $I(\nu, \theta)$, $\frac{\partial \tau(\nu, \theta, p)}{\partial \ln p}$ and $B(\nu, T(p)) \frac{\partial \tau(\nu, \theta, p)}{\partial \ln p}$ do not incorporate effects due to the instrument observational process. In a measured signal spectrum $S(\nu, \theta)$, the upwelling radiance within a narrow frequency interval $\delta \nu'$, which is given by $I(\nu', \theta) \delta \nu'$, gets redistributed by the

instrumental function $f_{inst}(\nu - \nu')$. Therefore, the measured signal can be represented by the well known convolution integral

$$S(\nu, \theta) = \int_0^\infty f_{inst}(\nu - \nu') I(\nu', \theta) d\nu'. \quad (2.9)$$

And the limits of integration can usually be replaced by cutoff frequencies associated with a bandpass filter. Analogously, the measurement process may be reflected in the weighting and energy contribution functions by the formulations

$$W(\nu, \theta, p) = \frac{\int_0^\infty f_{inst}(\nu - \nu') \frac{\partial \tau(\nu', \theta, p)}{\partial \ln p} d\nu'}{\int_0^\infty f_{inst}(\nu - \nu') d\nu'} \quad (2.10)$$

and

$$C(\nu, \theta, p) = \frac{\int_0^\infty f_{inst}(\nu - \nu') B(\nu', T(p)) \frac{\partial \tau(\nu', \theta, p)}{\partial \ln p} d\nu'}{\int_0^\infty f_{inst}(\nu - \nu') d\nu'}, \quad (2.11)$$

respectively. The denominators in equations 2.10 and 2.11 serve to normalize the functions with respect to instrument transmission. It is often useful to express the measured signal also normalized to instrument transmission, as

$$S_N(\nu, \theta) = \frac{\int_0^\infty f_{inst}(\nu - \nu') I(\nu', \theta) d\nu'}{\int_0^\infty f_{inst}(\nu - \nu') d\nu'}, \quad (2.12)$$

to enable direct comparison with atmospheric radiance or signals from other instruments in the same spectral region. Equations 2.9–2.12 can be used in an iterative process varying the instrumental function to establish the optimum instrument type as well as its associated parameters for the particular measurement of interest.

CHAPTER III

APPLICATION OF FABRY-PEROT INTERFEROMETRY

3.1 Introduction

The periodic nature of the Fabry-Perot instrument function can be advantageous when observation of periodic spectra is desired (Houghton, 1961; Wang, 1990). However, for most applications, additional optical elements are necessary to reduce the effect of unwanted passbands. This is frequently accomplished using additional Fabry-Perot etalons (Cocks, et al., 1980; Skinner et al., 1987) in conjunction with a bandpass filter. Appendix A contains an overview of Fabry-Perot interferometry theoretical concepts fundamental to this research application, including a procedure to formulate the associated instrument function including all significant broadening mechanisms. This chapter will introduce the concept of a multiple etalon configuration with particular emphasis placed upon the implementation of a double-etalon design. In addition, this chapter will contain a discussion of focal plane fringe detection techniques including a detailed summary of the new Circle to Line Interferometer Optical (CLIO) system (Hays, 1990; U.S. patent #4893003, 1990) technology.

3.2 Spectral Isolation Using Multiple Etalons in Series

A single order measurement of high spectral resolution is a hard task to achieve using an individual Fabry-Perot etalon. Application of a single etalon is restricted by the free spectral range ν_{fsr} of the interferometer; with the overall effective finesse being N_E , a single etalon FPI can observe at most N_E spectral elements of spectral width $\frac{\nu_{fsr}}{N_E}$. The transmission passbands become narrower as the etalon plate spacing is increased (ν_{fsr} decreased) or as the plate reflectivity is increased, but this also causes the undesirable neighboring passbands to move closer together and decreases the instrument throughput, respectively. However, two or more etalons may be used together in a series configuration to narrow the passband, and/or extend the effective free spectral range, and/or increase the contrast, depending upon the relationships between plate spacings and reflectivity coatings of the different etalons. Therefore, with the proper selection of etalon parameters, a multiple-etalon configuration can eliminate unwanted passbands and improve sideband rejection; and thus enable a successful single order, high spectral resolution measurement.

Application of two etalons in series can be traced back to 1897 (Hernandez, 1986) in work done by the pioneers of the FPI instrumentation, Charles Fabry and Alfred Perot; however, this was not for spectroscopic work. Use of multiple etalons for spectroscopic applications occurred somewhat later in time by other researchers; for example, in 1927 Houston (1927) successfully used two etalons in series. And since then, multiple etalons have been employed to handle many tasks. A double etalon in series configuration has been utilized by many researchers to distinguish closely spaced emission lines (i.e., Bradley and Kuhn, 1948; Daehler and Roesler, 1968; Daehler, 1970; Cannell and Benedek, 1970). Double etalon systems have been used

to record atomic spectral line emissions of $O(^1D)$ at $\lambda = 630$ nm in the daytime upper atmosphere from the earth's surface (Bens et al., 1965; Cocks et al., 1980). Triple etalon systems have also had many successful applications. The PEPSIOS (Poly-Etalon Pressure Scanning Interferometric Optical Spectrometer) device (Mack et al., 1963) has been the most popular of the triple etalon configurations; it consists of a pressure scanned triple etalon system arranged in a series configuration, having nearly equal etalon plate spacings. The PEPSIOS instrument has been applied to measurement of the 630nm dayglow (Barmore, 1977) as well as for many other atmospheric investigations (Hernandez, 1986; Vaughan, 1989). Rees et al. (1982) used a triple etalon configuration different from the PEPSIOS design to measure atmospheric O_2 absorption lines from a balloon platform. A successful satellite-based application of a multiple etalon has been achieved by the High Resolution Doppler Imager (HRDI) instrument (Hays, 1982; Skinner et al., 1987); it is a triple etalon Fabry-Perot interferometer designed to measure upper atmospheric winds from the Upper Atmosphere Research Satellite (UARS).

Since a double etalon configuration is the simplest representation of a multiple etalon system, and since such a configuration is intended for implementation in this research, its properties will now be considered in more detail. As mentioned earlier, various combinations of etalon parameters may be desired depending on what is expected from the multiple etalon system. Plate separation (or spacer) ratios can be grouped roughly into three categories (Roesler, 1974): multiple ratios corresponding to spacers having quite different lengths ($\frac{t_1}{t_2} > 2$), vernier ratios in which the spacers have approximately the same length ($\frac{t_1}{t_2} \approx 1$ or $\frac{t_1}{t_2} < 2$), or half-multiple ratios with $\frac{t_1}{t_2} = n + \frac{1}{2}$, where n is an integer. The etalon with the largest plate separation (or smallest ν_{fsr}) is usually referred to as the resolving etalon or high-resolution etalon

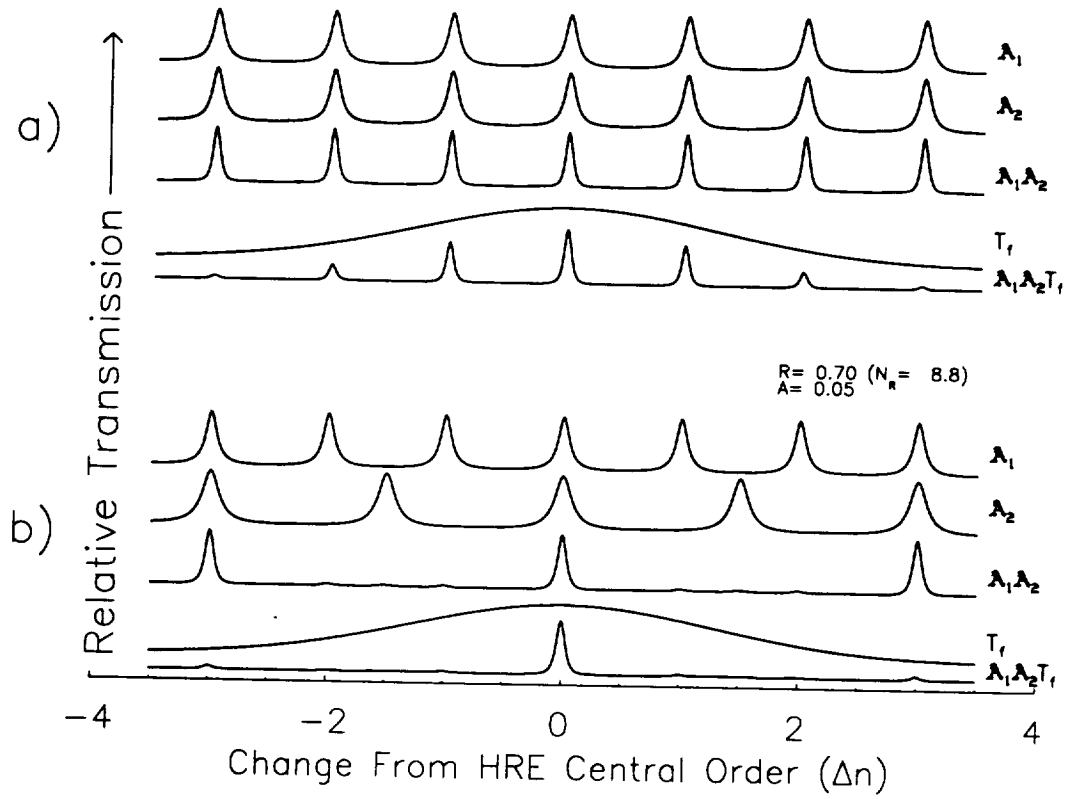


Figure 3.1: Transmission functions corresponding to two different double-etalon configuration scenarios. a) shows a case of equal plate spacings ($t_1/t_2 = 1$). b) corresponds to a case of etalon plate spacers having a half-multiple ratio with $t_1/t_2 = 1.5$.

(HRE), as the width of its passbands pretty much determine the multiple etalon resolution. Conversely, the etalon with the smallest plate spacing (or largest ν_{fsr}) is commonly called the low-resolution etalon (LRE) or suppression etalon, with its purpose being to block some of the undesired or parasitic orders of the HRE and thus yield an extension of the effective free spectral range. Transmission functions resulting from a couple of different double-etalon configuration scenarios are illustrated in Figure 3.1. Figure 3.1a) shows a case of equal plate spacings ($\frac{t_1}{t_2} = 1$). Ignoring coupling between etalons and broadening effects, the multiple etalon transmission function can be given approximately by the product of the individual Airy functions ($E = \mathcal{A}_1 \cdot \mathcal{A}_2 \cdots \mathcal{A}_m$); as shown in the Figure, for a double-etalon configuration this yields a transmission function of higher resolution and contrast. The broad wing characteristic of a single-etalon transmission function is greatly suppressed using another etalon in the vernier spacer ratio case, enhancing the purity of spectral isolation. However, using equal plate spacings (Figure 3.1a)) will not extend the free spectral range through blockage of passbands. A Gaussian function is assumed for the shape of the ultra-narrow bandpass filter. As depicted in Figure 3.1a), a multi-order transmission function results from the double-etalon/filter system, since the bandpass filter width (HWHM) is greater than the effective free spectral range of the double etalon. Figure 3.1b) corresponds to a case of etalon plate spacers having a half-multiple ratio with $\frac{t_1}{t_2} = 1.5$. Extension of the effective free spectral range is achieved for this case, as is obvious from the double-etalon transmission function plot. However, due to the finite contrast of etalons, small parasitic (or ghost) peaks result where the blocked passbands once existed. Incorporating the bandpass filter further reduces the magnitude of the parasitic peaks and nearly eliminates unwanted passband orders; thus yielding a single-order transmission function for the double-

etalon/filter configuration. It should be noted that various pairs of spacer ratios can produce a transmission function with the same extension of free spectral range, having ghost peaks of approximately the same strength but in possibly different locations.

The method of coupling Fabry-Perot etalons in series can be quite simple. Etalons having the same working aperture diameter may be placed in a series with their axes parallel without use of any intermediate optics. The etendue is limited by the etalon with the largest plate separation. Alternatively, etalons of different diameters may be coupled using magnification between the etalons without loss of etendue (Vaughan, 1989). As demonstrated by the Airy function (equation A.1), the angular positioning (in the fringe plane) of a transmitted wavenumber does not depend upon etalon plate separation; this implies that the wavenumber dependence on angle for a multiple-etalon system is the same as that for a single etalon, when the multiple-etalon configuration consists of a simple series without employing intermediate magnification. Therefore spectral isolation apertures for a simple series polyetalon can be chosen in the same way as for a single etalon. Conversely, etalons of different diameter coupled with intermediate magnification each have a different wavenumber dependence on angle, since the angle subtended at each etalon by a point on the intermediate aperture is different for each etalon (Roesler, 1974). Consequently, only a single-channel measurement can be made at the central fringe, where the etalons of different plate spacing are "in register."

As mentioned earlier, the transmission of a simple series of etalons can be given by the product of the individual Airy functions, ignoring reflections between etalons and broadening effects. Broadening alone can be incorporated by replacing the Airy functions with the plate defect broadened Airy functions, and then convolving the

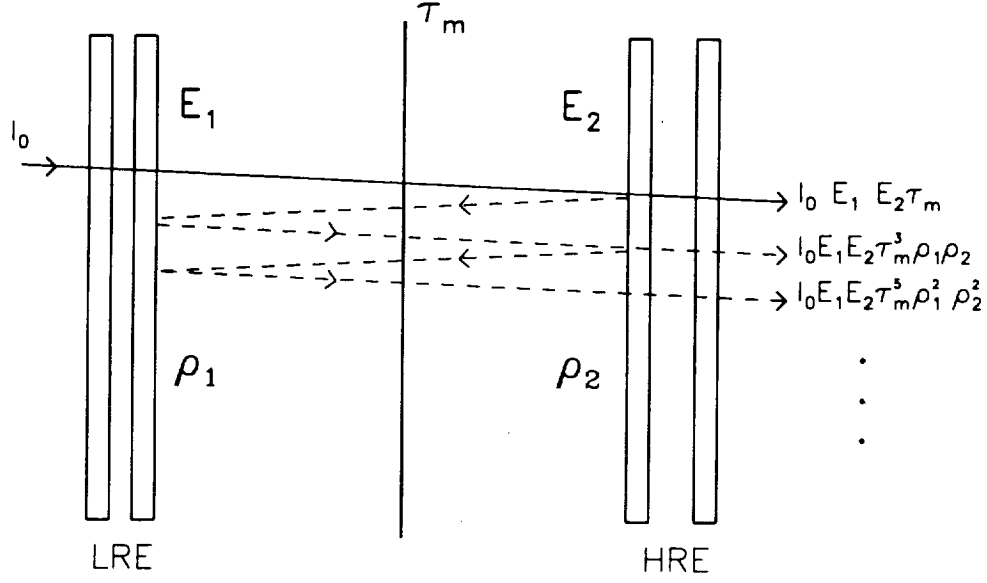


Figure 3.2: A simple double-etalon series configuration. The transmission through the attenuating medium is defined as τ_m , while the etalon transmittances and reflectances are denoted by the E_i and ρ_i , respectively.

aperture function L with the resulting product (i.e., $E = L * [(\mathcal{A}_1 * D_1) \cdot (\mathcal{A}_2 * D_2)]$, for a double-etalon configuration ignoring reflections). The inter-etalon reflections can be handled in several ways: they can simply be ignored if their effect is shown to be insignificant; an attempt can be made to eliminate them by introducing a slight tilt of etalon axes with respect to each other, which can interfere with alignment for a multichannel configuration, or by using a system of polarizers (Hernandez, 1986) at the cost of severely reducing throughput; alternatively, their impact can be reduced to acceptable levels by placement of an attenuating medium in between the etalons. This last method involving placement of a loss factor in between the etalons can be the simplest and most effective approach (Vaughan, 1989).

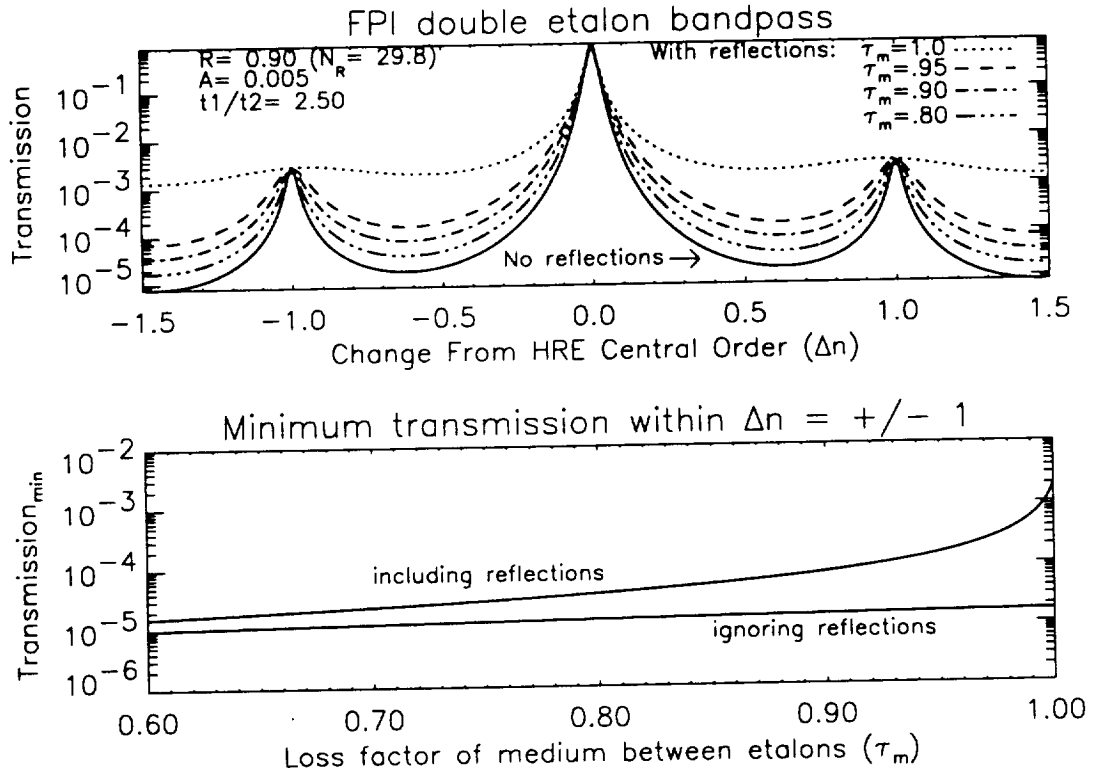


Figure 3.3: Transmission through a double-etalon system incorporating inter-etalon reflections and with an attenuating medium placed in between the etalons. Defect broadening effects are not included. a) shows transmittance through the system including reflections for several loss values of the intermediate medium. The solid curve represents the transmission with reflections ignored and without any intermediate attenuation ($\tau_m = 1$). b) illustrates the minimum achievable transmittance within one order of the HRE central order as a function of loss factor between the etalons, with and without the inclusion of inter-etalon reflections.

Mack et al. (1963) derived a transmittance function for a triple etalon system possessing attenuating media between adjacent etalons while incorporating reflections between etalons. Their effort was associated with the PEPSIOS pilot model instrument, however, their resulting transmission expression is applicable to other triple-etalon systems (Skinner et al., 1987). An analogous expression can be derived for a double-etalon system. First one can consider a simple double-etalon series configuration as illustrated in Figure 3.2. The transmission through the attenuating medium is defined as τ_m , whereas the etalon transmittances and reflectances are denoted by the E_i and ρ_i , respectively. The etalon transmittances may be expressed by etalon transmission functions (equation A.17) as $E_i = \mathcal{A}_i * D_i$, and the etalon reflectivities can be defined simply by $\rho_i = 1 - E_i$ (Born and Wolf, 1975; Skinner et al., 1987; Vaughan, 1989). This simple relationship serves as a very good approximation since absorption is small in dielectric reflecting film coatings, leaving scattering principally responsible for reduction of intensity; and the scattered ray travels very close to the reflected and transmitted rays (Skinner et al., 1987). The expressions to the right of the HRE in Figure 3.2 show the contributions to the total intensity of transmitted radiation from direct transmission and from the successive reflections between the etalons. With an initial intensity of radiation I_0 entering the LRE, the total intensity of radiation exiting the double etalon system is obtained by summing the direct transmission and successively reflected components, and can be given by

$$I = I_0 E_1 E_2 \tau_m (1 + \tau_m^2 \rho_1 \rho_2 + \tau_m^4 \rho_1^2 \rho_2^2 + \cdots). \quad (3.1)$$

Therefore, the system transmittance would be $E_D = \frac{I}{I_0}$, or

$$E_D = E_1 E_2 \tau_m [1 + \sum_{i=1}^{\infty} (\tau_m^2 \rho_1 \rho_2)^i] = E_1 E_2 \tau_m \sum_{i=0}^{\infty} (\tau_m^2 \rho_1 \rho_2)^i. \quad (3.2)$$

The Maclaurin expansion of $\frac{1}{1-z}$ is the following geometric series: $\frac{1}{1-z} = \sum_{n=0}^{\infty} Z^n$.

Utilizing this series expression, the double-etalon transmission function can be expressed as

$$E_D = \frac{E_1 E_2 \tau_m}{1 - \tau_m^2 \rho_1 \rho_2}. \quad (3.3)$$

Figure 3.3 shows how inter-etalon reflections affect a double-etalon system transmission function for different values of attenuation for an intermediate medium τ_m . As τ_m is reduced from 1.0 to .95, the passband wing suppression is greatly improved with the minimum system transmittance dropping by more than one order of magnitude; thus illustrating how intermediate placement of a low attenuation medium can improve the degrading transmission effects introduced by incoherent reflections between etalons.

The effect of a finite sized aperture can be accounted for in a double-etalon system similar to its incorporation into a single-etalon system, and analogous to equation A.21 the double-etalon FPI instrumental transmission function can be represented by

$$T_D(\Psi) = E_D(\Psi) * L(\Psi). \quad (3.4)$$

The construction of a double-etalon FPI instrumental transfer function including broadening effects and incoherent reflections between etalons is summarized schematically in Figure 3.4. Figure 3.5 qualitatively depicts the formation of a double-etalon instrumental transfer function from the Airy (\mathcal{A}), defect (D), aperture (L), and bandpass filter (T_F) functions.

It was mentioned earlier (Section A.3) that for optimum FPI transmittance with minimal loss of finesse (or for optimum luminosity-resolution product (LRP)) the reflectivity finesse should be matched to the defect finesse of the plates ($N_R \approx N_D$). However, the requirements are more severe for a poly-etalon system (Hernandez, 1986; Mack et al., 1963; Roesler, 1969, 1974). Plate defect tolerances are smaller

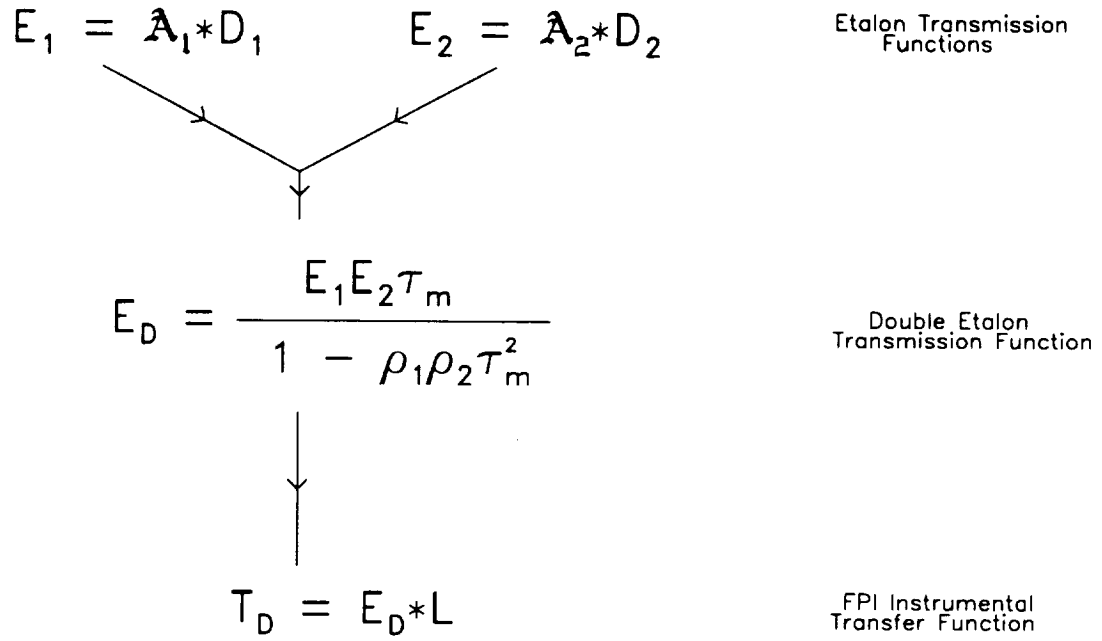


Figure 3.4: Schematic representation of the construction of a double-etalon FPI instrumental transfer function including broadening effects and incoherent reflections between etalons. The loss factor of an attenuating medium (τ_m) is included to reduce the impact of inter-etalon reflections.

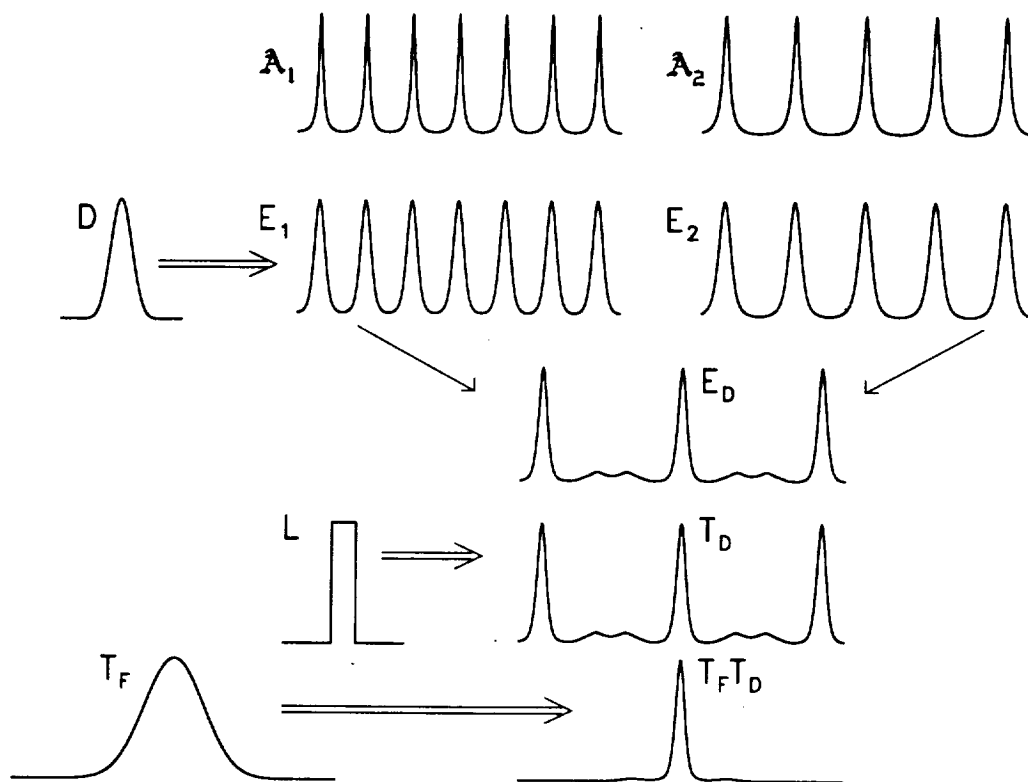


Figure 3.5: Qualitative diagram summarizing the formation of a double-etalon instrumental transfer function from the Airy (\mathcal{A}), defect (D), aperture (L), and bandpass filter (T_F) functions. The Airy functions simulated correspond to an etalon plate spacer ratio of $\frac{t_1}{t_2} = 1.5$. All plots shown except those for the defect and aperture functions correspond to normalized transmission functions having the same horizontal and vertical scales. However, the defect and aperture functions are plotted with an abscissa having one-seventh the range of the abscissa scale used in the other plots; this was done to better illustrate their functional forms.

for a series of etalons with fixed reflectivity finesse than those acceptable for a single etalon system in order to maintain optimum transmittance. Conversely, the plate reflectivities in a multiple etalon system with fixed defect finesse must be significantly lower than those for a single etalon to achieve the same peak transmittance (Roesler, 1969); and the impact of lower plate reflectivities is lower spectral purity (or resolution) for the multi-etalon configuration. A complete treatment for a series of etalons would involve specification of each individual defect function and relative orientations of the etalons. However, the analysis can be simplified by making some basic assumptions while maintaining the essential characteristics of the multi-etalon transmission function. This research will assume the same symmetric defect distribution function to be valid for each etalon ($D_1 = D_2 = D(\Psi)$), and that correlations among defects of the different etalons can be ignored. In addition, the etalons are assumed to be aligned perfectly with respect to one another.

3.3 Methods of Tuning and the CLIO System

The discussion in Section A.2 illustrated that the Airy function reaches a local maximum in transmission wherever $2\mu t\nu \cos \theta = n$, the integer order of interference. It is obvious from this equation that spectral information can be inferred through variation of μ , t , or θ . This forms the basis for the three conventional types of spectroscopic operation of a Fabry-Perot interferometer. The spectroscopic variable can be the index of refraction of the medium between the plates μ (pressure scanning), the spacing between the plates t (mechanical scanning), or the angle of incidence of the incident radiation θ (spatial scanning). A schematic representation of these spectroscopic variables was given in Figure A.1.

In pressure scanning, the tuning or scanning is accomplished by varying the

pressure within the etalon cavity, thus changing the refractive index of the medium between the plates μ . This is usually used in conjunction with a single detector centrally located within the fringe plane; thus allowing observation of a single spectral “channel” at any given time.

Mechanical scanning usually has one plate mechanically translated to vary the spacing between the plates t . As with pressure scanning, this method typically incorporates usage of a single centrally placed detector providing a single-channel measurement.

Scanning spatially in the fringe plane is the foundation for many “spatial scanning” techniques. This method is sometimes referred to as “angular scanning” since the angle θ changes with spatial position in the fringe plane, or as “tilt scanning” since θ can be changed by simply tilting the etalon. An aperture/detector can be scanned across fringes in the image plane; this would yield a single-channel measurement. Conversely, a photographic plate or imaging photodetector can be fixed in the fringe plane; this would constitute a multichannel measurement, since multiple spectral channels of information can be observed simultaneously. Alternatively, many detectors or a diode array may be positioned in the fringe plane to also yield a multichannel measurement capability.

The scanning of multiple Fabry-Perot interferometers is the same as for single etalons for pressure scanning and, in the case of a simple series configuration (without intermediate magnification), spatial scanning; since the tuning is independent of plate spacers. The spatial scanning technique is not applicable to etalons coupled with magnification since the angular dependence of transmittance is altered by the coupling. Tuning by mechanical scanning is dependent upon plate spacing; therefore, any technique to mechanically scan multiple etalons must enable separate adjustable

scanning mechanisms for each etalon.

The intended application of this research is for a satellite-based instrument remotely monitoring atmospheric ozone through utilization of a nadir-viewing geometry. Accordingly, a multichannel spectral measurement capability is required to enable any sort of ozone vertical distribution discrimination while maximizing signal-to-noise of the measurements; thus necessitating implementation of a spatially scanning configuration, from the above discussion of spectroscopic tuning techniques. Additionally, a satellite-based instrument further restricts the potential choices of tuning techniques (i.e., implementation of a photographic plate would not be feasible for a space-based application). Historically, many approaches have been pursued to develop multichannel optical devices. Several instrument designs made in the past have had the different spectral channels of information detected simultaneously at separate detectors. For example, Hirschberg and Platz (1965) employed a technique using concentric mirrors in the fringe plane (each with a slight tilt with respect to the optical axis) to reflect light onto separate photodetectors. However, recent advances in detector technology have led to achieving simultaneous multichannel operation using a single detector unit (Hernandez, 1986; Vaughan, 1989); for example, position sensitive devices such as the imaging photodiode (Rees et al., 1981) and more recently the image plane detector (Hays et al., 1981; Killeen et al., 1983), and a charge coupled device (CCD) as an imaging device (Abreu and Skinner, 1989). One of the most ingenious devices with demonstrated success for space-based applications, the image plane detector (IPD) developed for the Dynamics Explorer Fabry-Perot Interferometer (DE-FPI) satellite mission provided a multichannel spectral measurement through spatially scanning the interference fringe plane using shaped anodes on a single detector unit. The DE-IPD is an electro-optical device that converts energy

reaching the twelve concentric anodes in the image plane (which mimic the fringe pattern of the etalon) into a set of discrete electron pulses. This type of approach has some significant advantages over conventional scanning techniques (Killeen et al., 1983): the need for mechanical scanning is eliminated, and it provides a signal multiplex advantage by detecting all the energy in the imaged ring pattern. In addition, a longer dwell time per spectral channel may be achieved since all channels are measured simultaneously; this is of particular interest for satellite applications where the instrument field-of-view (FOV) is in motion relative to the target scene. The successful implementation of an IPD on the DE-FPI led to a similar multiple anode detector being developed for the High-Resolution Doppler Imager (HRDI) instrument aboard the Upper Atmosphere Research Satellite (UARS) (Skinner et al., 1987; Hays and Wang, 1991). However, this more advanced multichannel anode device developed for the HRDI instrument has thirty-two channels instead of the twelve used with the DE-FPI.

Image plane detectors have a proven track record in space-based applications, however, lagging detector technology at longer wavelengths prevents their development in the IR region of the spectrum. To circumvent this IR-FPI detection difficulty, this research will incorporate the new Circle to Line Interferometer Optical (CLIO) system (Hays, 1990; U.S. patent #4893003, 1990) technique to accomplish the spatial scanning of the fringe plane, and thus enable multichannel operation. A successful laboratory demonstration of this technique has been carried out in the Space Physics Research Laboratory (SPRL) at the University of Michigan (Hays, 1990). In addition, the CLIO system has been included as a vital ingredient in the Multi-Order Etalon Sounder (MOES) instrument design concept (Wang, 1990).

The CLIO system converts circular fringes of an FPI into a linear pattern which

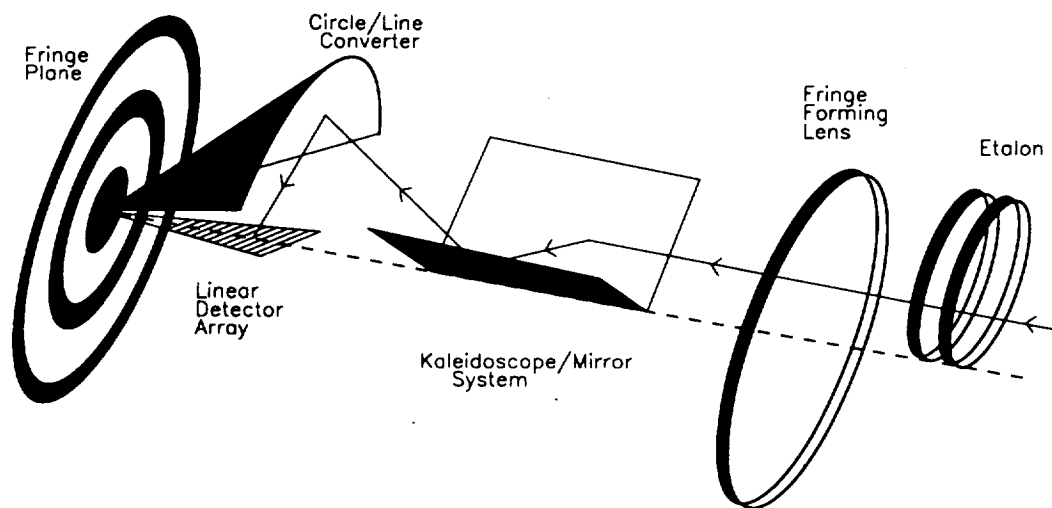


Figure 3.6: Conceptual diagram of the CLIO system optical configuration used to convert a circular Fabry-Perot fringe pattern into a linear pattern along the axis of a 45 degree half-angle internally reflecting transforming cone.

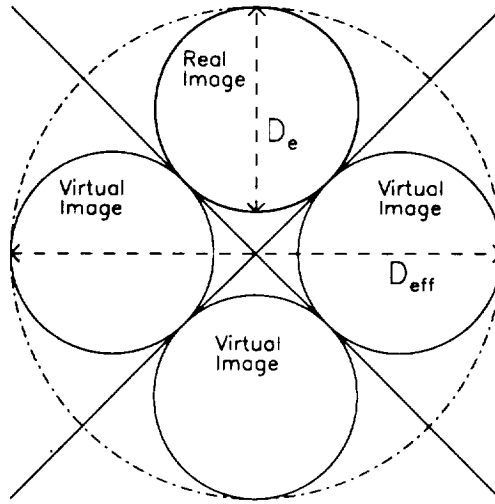


Figure 3.7: View from fringe plane of the CLIO system without the circle to line converter in place.

can be detected by a conventional linear array detector; thus resembling a standard spectrometer output, while enabling utilization of the high throughput of an FPI. In spectral regions where an IPD can be utilized, application of a CLIO system may be more advantageous; the CLIO system can put the same amount of energy on a smaller detector surface area thus improving the signal-to-noise, and implementation of a CLIO system should be more cost beneficial than the corresponding IPD (Hays, 1990). The circle to line transformation is accomplished with a 45 degree half-angle internally reflecting cone segment of 90 degree azimuthal extent as the primary optical element; it is centered above the detector plane while its vertex is centered in the fringe plane. In addition, an optical kaleidoscope is located in the optical path prior to the cone to enable transformation of the entire circular ring pattern (instead of just one-fourth of the ring energy being transformed by the 90 degree cone segment alone). The basic optical configuration of the CLIO system is illustrated in Figure 3.6. The mirrored kaleidoscope concentrates the entire fringe pattern onto the 90 degree cone sector. The cone then reflects the incident energy onto the detector array centered on the cone axis below. The overall effect of the kaleidoscope and 45 degree half-angle cone combination is to transform energy that would have intersected a circle of radius r_0 in the fringe plane onto a detector element whose center intersects the cone axis at a distance r_0 from the cone vertex. The kaleidoscope converts all of the energy that would be incident on a complete circle in the fringe plane into a small angular segment, corresponding to the azimuthal extent of the cone segment. For a 90 degree cone sector, this can be accomplished using a simple 90 degree “telekaleidoscope” (Hays, 1990): this optical device consists of a segment of a telescope objective enclosed by two mirrors joined at the optical axis of the lens system and terminated with the transforming cone segment. If viewed from fringe

plane without the cone present, it would appear as if there were four identical etalons and fringe forming lenses; the real image would be above three virtual images in a crossed pattern (see Figure 3.7). Thus implying an etalon “effective” diameter D_{eff} greater than the actual etalon diameter D_e . From the simple geometrical relationship depicted in Figure 3.7 it can easily be shown that

$$D_{eff} = D_e(1 + \sqrt{2}) \approx 2.414D_e. \quad (3.5)$$

This expression will be of use when design considerations are addressed in the next chapter.

CHAPTER IV

DOUBLE-ETALON INSTRUMENT DESIGN TO ACHIEVE SINGLE-ORDER HIGH-RESOLUTION MEASUREMENT OF ATMOSPHERIC OZONE

4.1 Introduction

The water vapor continuum absorption provides the most significant impact of all interfering species within the 10 μm spectral region, since its effect cannot be minimized through channel spectral positioning as can be done for spectral lines of interfering species. In addition, water vapor is most significant for atmospheres within the tropics, geographical regions for which tropospheric ozone measurements are highly desirable. Consequently, this chapter will begin with a discussion of the water vapor continuum absorption present in the 10 μm spectral region. A method for calculating the upwelling radiance reaching a satellite sensor, including ozone and all significant interfering species, will then be presented. Simulations assessing the radiance sensitivity to atmospheric ozone perturbations will be shown to suggest the necessary instrument spectral resolution and approximate placement of the measurement band. Double-etalon design specifications to achieve the desired spectral isolation will then be developed. Demonstration of ozone information content of the measured signals will follow a justification for spectral placement of the instrument

band and corresponding channels. The chapter will then conclude with addressing additional design considerations leading up to a signal-to-noise ratio analysis.

4.2 Water Vapor Continuum Absorption in the 10 μm Spectral Region

Absorption by water vapor (H_2O) in the 800-1200 cm^{-1} spectral window region of the lower atmosphere is dominated by continuum absorption. An accurate calculation of the absorption in this spectral region must include the contribution from continuum absorption as well as the local line absorption. However, the continuum absorption remains poorly understood and difficult to accurately model due to the inability of current theories to adequately explain existing observations, the large scatter present among laboratory measurements, and the uncertainty associated with extrapolating limited laboratory observations to the pressure and temperature ranges present in the open atmosphere. Nevertheless, many measurements of water vapor absorption in the thermal IR atmospheric window have been performed over the years by a variety of instrumentation techniques (Grant, 1990), from which empirical water vapor continuum absorption models can be deduced.

Water vapor continuum absorption is spectrally unstructured absorption by H_2O vapor, excluding calculable contributions from the wings of lines within some defined proximity. However, different cutoff distances for line wings have been suggested in the literature, ranging from 1 cm^{-1} (Burch and Alt, 1984) to 25 cm^{-1} (Clough et al., 1980) with several values in between. Although these undesirable differences in wing cutoff definitions may cause some confusion, they are probably insignificant relative to the laboratory measurement uncertainties (Grant, 1990). Two main theories have been developed to explain the poorly understood water vapor continuum absorption:

1) The water vapor absorption line shape model, where the continuum is assumed to be caused by extreme wings of absorption lines of H_2O molecules (monomers); and 2) the water vapor cluster model, where clusters consisting of multiple H_2O molecules bound together are assumed to result in the observed absorption continuum. The water vapor cluster model is also known as the water dimer model since the fraction of and corresponding contribution by water dimer dominate over those of all the other water clusters.

Five different theoretical line shapes are conventionally used for collision-broadened water vapor absorption lines (Burch, 1981); based upon different sets of assumptions regarding the nature of collisions, they all accurately represent H_2O absorption near line centers and degrade to inadequate representation toward the far wings. In fact, none of the theoretical line shapes can accurately predict the extreme wing absorption without application of an empirically derived line shape correction. Some characteristics of the continuum not well explainable by monomer theory, such as the pressure square dependence and negative temperature dependence of absorption, have supported the water vapor dimer explanation of the observed continuum absorption (Suck et al., 1982). However, current H_2O dimer theories fail to support the water dimer as being the only or even dominant contributor to this continuum absorption in the IR region; accordingly, the corresponding models fail to explain the spectral character of the absorption. Therefore it is common for one to assume that water vapor continuum absorption is caused by absorption from the wings of distant strong lines (Thomas and Nordstrom, 1985), which are significant due to collisions among water vapor molecules (self-broadening) and collisions between H_2O vapor and other species (foreign-broadening); observations show that the absorption coefficient increases with the water vapor partial pressure (self-broadening) and with

the total pressure (foreign-broadening). Prior to knowledge that different broadening is caused by different molecules, the combined broadening effect was considered as pressure broadening. However, water vapor self-broadening is now known to usually be much stronger than foreign-broadening.

Considering only water vapor absorption, the absorption coefficient can be expressed as

$$\kappa = \kappa_l(\nu, T, p) + \kappa_c(\nu, T, p), \quad (4.1)$$

where κ_l is the water vapor line absorption coefficient (as discussed in section 2.4) which can be accurately computed using the line-by-line method (see section 4.3), and κ_c is the water vapor continuum absorption coefficient.

As a commonly implemented first-order approximation, κ_c can be expressed as a linear function of the water vapor partial pressure and the partial pressure of foreign broadening molecules,

$$\kappa_c(\nu, T, p) = C_s(\nu, T)p_{H_2O} + \sum_i C_{f_i}(\nu, T)p_{f_i}, \quad (4.2)$$

where C_s and C_{f_i} are the coefficients representing self-broadening and foreign-broadening due to the i th species, respectively, in units of $\text{cm}^2 \text{ molecules}^{-1} \text{ atm}^{-1}$; p_{H_2O} and p_{f_i} are the water vapor partial pressure and partial pressure of the i th broadening component, respectively, in units of atmosphere. Since dry air is composed primarily of molecular nitrogen (N_2) and molecular oxygen (O_2), the foreign-gas broadening can be accurately estimated by considering only N_2 and O_2 . Equation 4.2 can then be expressed as

$$\kappa_c(\nu, T, p) = C_s(\nu, T)p_{H_2O} + C_{N_2}(\nu, T)p_{N_2} + C_{O_2}(\nu, T)p_{O_2}, \quad (4.3)$$

or by

$$\kappa_c(\nu, T, p) = C_s(\nu, T)p_{H_2O} + C_{air}(\nu, T)[p - p_{H_2O}]; \quad (4.4)$$

p is the total atmospheric pressure, so that $p - p_{H_2O}$ is the partial pressure of dry air; C_{air} is the foreign broadening coefficient for dry air, which, upon considering the proportions of N_2 and O_2 in dry air, can be given by

$$C_{air} = C_{N_2} [0.79 + 0.21 \frac{C_{O_2}}{C_{N_2}}]. \quad (4.5)$$

The self- and foreign-broadening coefficients derived by Zhao (1992) for the 500-1200 μm region will be utilized for modeling the water vapor continuum absorption in this research. Zhao (1992) applied linear regression to the Burch and Alt (1984) data using boundary values for good connection to nearby spectral regions in the fitting process; in addition, LOWTRAN6 (1020 to 1200 cm^{-1}) data were used in deducing the foreign-broadening coefficients due to lack of Burch and Alt (1984) data. Zhao (1992) arrived upon the following set of expressions:

$$C_{s_{284}}(\nu) = [15218 \exp(-0.004216\nu) - 20.42] \times 10^{-24}, \quad (4.6)$$

and

$$C_{s_{296}}(\nu) = [34003 \exp(-0.005884\nu) + 47.56] \times 10^{-24}, \quad (4.7)$$

representing the self-broadening coefficients (in units of $molecules^{-1} cm^2 atm^{-1}$) for 284 K and 296 K, respectively, as a function of wavenumber ν (cm^{-1}); and accounting for the self-broadening coefficient temperature dependence,

$$C_s(\nu, T) = C_{s_{296}}(\nu) \left[\frac{C_{s_{296}}(\nu)}{C_{s_{284}}(\nu)} \right]^{(23.67 - \frac{7005.3}{T})}, \quad (4.8)$$

the nitrogen foreign-broadening coefficient wavenumber dependence at a temperature of 296 K can be given by

$$C_{N_{296}}(\nu) = [1.3167 \times 10^{-6} \exp(.01472\nu) + 16.338 + 1.6219 \times 10^6 \exp(-.012383\nu)] \times 10^{-26}; \quad (4.9)$$

and assuming the temperature dependence of the foreign-broadening coefficient to be the same as that for the self-broadening coefficient at a particular wavenumber, gives

$$C_N(\nu, T) = C_{N_{296}}(\nu) \left[\frac{C_{s_{296}}(\nu)}{C_{s_{284}}(\nu)} \right]^{(23.67 - \frac{7005.3}{T})}; \quad (4.10)$$

assuming the ratio of broadening coefficients to be the same as that for broadening efficiencies and that the O_2 to N_2 broadening efficiency ratio in the 10 μm region is approximately equal to 0.75, equation 4.5 can be used to represent air broadening by

$$C_{\text{air}}(\nu, T) = 0.95 C_N(\nu, T); \quad (4.11)$$

in addition, the self- and foreign-broadening coefficients are assumed to be constant for temperatures lower than 273 K, that is

$$C_s(\nu, T) = C_s(\nu, T = 273 \text{ K}) \quad (4.12)$$

and

$$C_N(\nu, T) = C_N(\nu, T = 273 \text{ K}), \quad (4.13)$$

when $T < 273 \text{ K}$.

4.3 Atmospheric and Total Upwelling Radiance Calculations

The expression for upwelling radiance given in equation 2.1 of section 2.4 can be reformulated into a more direct form which lends itself to higher computational accuracy as

$$I(\nu, \theta) = \epsilon_s(\nu, \theta) B(\nu, T_s) \tau(\nu, \theta, p_s) + \int_{\tau(\nu, \theta, p_s)}^1 B(\nu, T(p)) d\tau(\nu, \theta, p), \quad (4.14)$$

where the integral over transmittance corresponds to along the path of radiation. Simulations for this research will be restricted to a nadir (zenith angle $\theta = 0^\circ$) view-

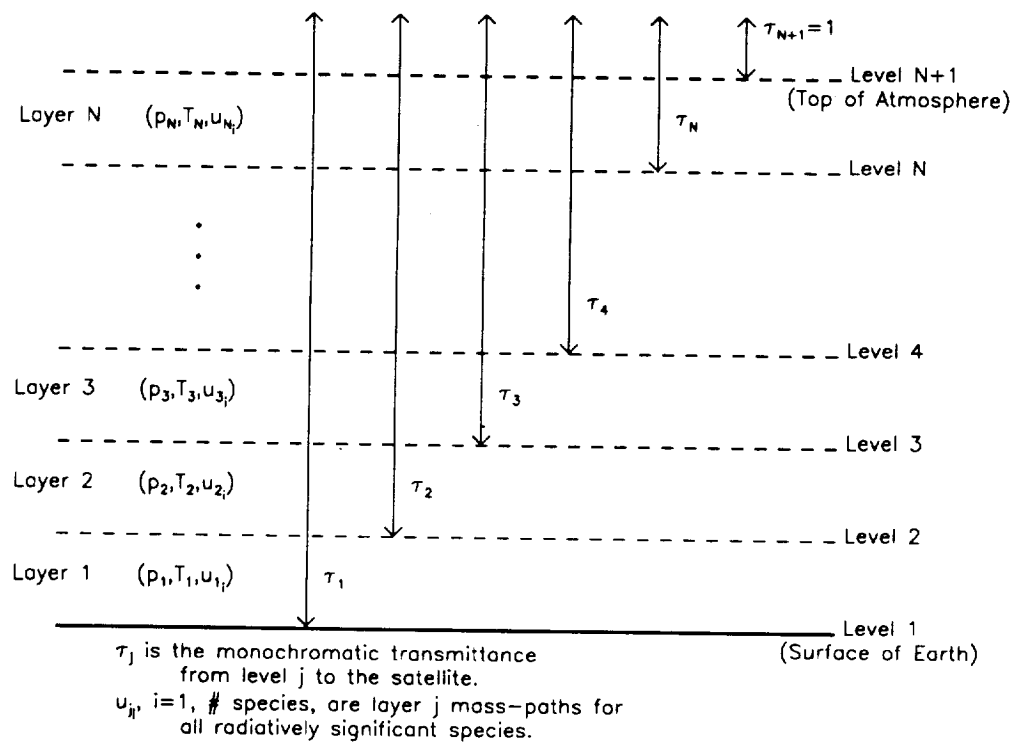


Figure 4.1: Representation of an inhomogeneous atmosphere by a sequence of homogeneous slabs.

Level	p (mb)	z (km)	Level	p (mb)	z (km)
1	1013.00000	0.00	24	65.69215	18.90
2	954.19305	0.50	25	50.34081	20.60
3	887.83746	1.10	26	38.63487	22.30
4	824.81378	1.70	27	29.26718	24.10
5	765.61346	2.30	28	21.56443	26.10
6	710.05902	2.90	29	15.22454	28.40
7	657.54083	3.50	30	9.71346	31.30
8	608.53101	4.10	31	6.73928	33.80
9	562.28705	4.70	32	5.17779	35.80
10	519.03278	5.30	33	3.97051	37.80
11	478.62231	5.90	34	3.04534	39.60
12	440.59213	6.50	35	2.32149	41.60
13	405.28329	7.10	36	1.78688	43.60
14	372.07111	7.70	37	1.36579	45.70
15	341.19833	8.30	38	0.54675	53.00
16	312.53714	8.90	39	0.38222	55.80
17	285.69214	9.50	40	0.28175	58.10
18	260.92990	10.10	41	0.20424	60.50
19	237.78906	10.70	42	0.14409	63.00
20	182.18594	12.40	43	0.10126	65.50
21	141.70000	14.00	44	0.06905	68.10
22	110.21021	15.60	45	0.05220	70.00
23	85.76624	17.20	46	0.00032	100.00

Table 4.1: Stratification used to represent the 1976 U.S. Standard atmosphere.

ing geometry. Since off-nadir views, although usually necessary to achieve the desired global coverage using a small horizontal footprint, will not add any further insight into this instrument feasibility analysis; this removes the zenith angle dependence of the radiative transfer equation (RTE). In addition, the definite integral term present in the RTE is usually estimated by a numerical approximation (or quadrature) technique. Dropping the zenith angle dependence and utilizing the trapezoidal rule for integral approximation, equation 4.14 can be expressed as

$$I(\nu) \approx \epsilon_s(\nu)B(\nu, T_s)\tau(\nu, p_s) + \sum_{i=1}^{\#layers} B_i(\nu, T(p))\delta\tau_i(\nu, p). \quad (4.15)$$

The \approx sign is shown since implementation of a quadrature technique provides an estimate of the definite integral, the accuracy of which depends directly on the number of subintervals used in the evaluation; in theory, the quadrature induced error $\rightarrow 0$ as the number of subintervals $\rightarrow \infty$. Radiation trajectories through the non-homogeneous (real) atmosphere are usually treated by dividing the atmosphere into many vertically successive homogeneous layers (i.e., layers for which the pressure and temperature can be represented as constants). The monochromatic transmittance through the atmosphere can then be obtained from the product of the transmittances for each of the homogeneous layers along the path. The number of layers must be \geq the number of levels for which atmospheric parameter estimation is desired, and it must be large enough to achieve sufficient accuracy in the radiance calculations (i.e., an iterative process whereby the number of layers is successively increased until convergence of the calculated radiance is achieved can be used). In general, finer atmospheric stratification produces higher accuracy in the radiance calculations at the cost of computational speed. A schematic diagram illustrating the representation of an inhomogeneous atmosphere by a sequence of homogeneous slabs is shown in Figure 4.1. For the N layer atmosphere depicted in Figure 4.1, equation 4.15 can be

expressed as

$$I(\nu) \approx \epsilon_s(\nu)B(\nu, T_s)\tau_1(\nu, p_s) + \sum_{i=1}^N B(\nu, T_i)[\tau_{i+1}(\nu) - \tau_i(\nu)], \quad (4.16)$$

where τ_j is the monochromatic transmittance from level j to the satellite. A 45 layer atmosphere was used to represent the real atmosphere in calculations for this study. The exact stratification varied slightly depending on the particular atmosphere's thermal characteristics; Table 4.1 shows the layering used for the 1976 U.S. Standard Atmosphere (Anderson et al., 1986) which was used in the majority of calculations.

The atmospheric parameter retrieval process for remote sounding applications usually requires the upwelling radiance to be accurately calculated at high spectral resolution. Accordingly, calculations of atmospheric transmittances must also be made with high accuracy and at high spectral resolution. This is customarily accomplished using a line-by-line (LBL) method for transmittance calculations (Drayson, 1966), whereby the contribution to the absorption coefficient from all significant spectral lines is accounted for at each frequency of interest. At the expense of computational speed, the numerical accuracy of the LBL method can be made as high as desired, making the uncertainty of spectral line parameters (and their associated atmospheric variations) the limiting factor in the overall transmittance calculation accuracy. LBL calculations for this research were performed using a slightly modified (for adapting to this particular research application) computer program developed by Professor S. Roland Drayson at the Space Physics Research Laboratory (SPRL) of the University of Michigan; this LBL program originates from earlier work by Drayson (1966). The necessary atmospheric parameters (pressures p_i , temperatures T_i , and masspaths u_i) required to uniquely describe the vertical sequence of horizontally homogeneous slabs were calculated using a modified version of the air mass computer program FSCATM (Gallery et al., 1983) contained within the Fast At-

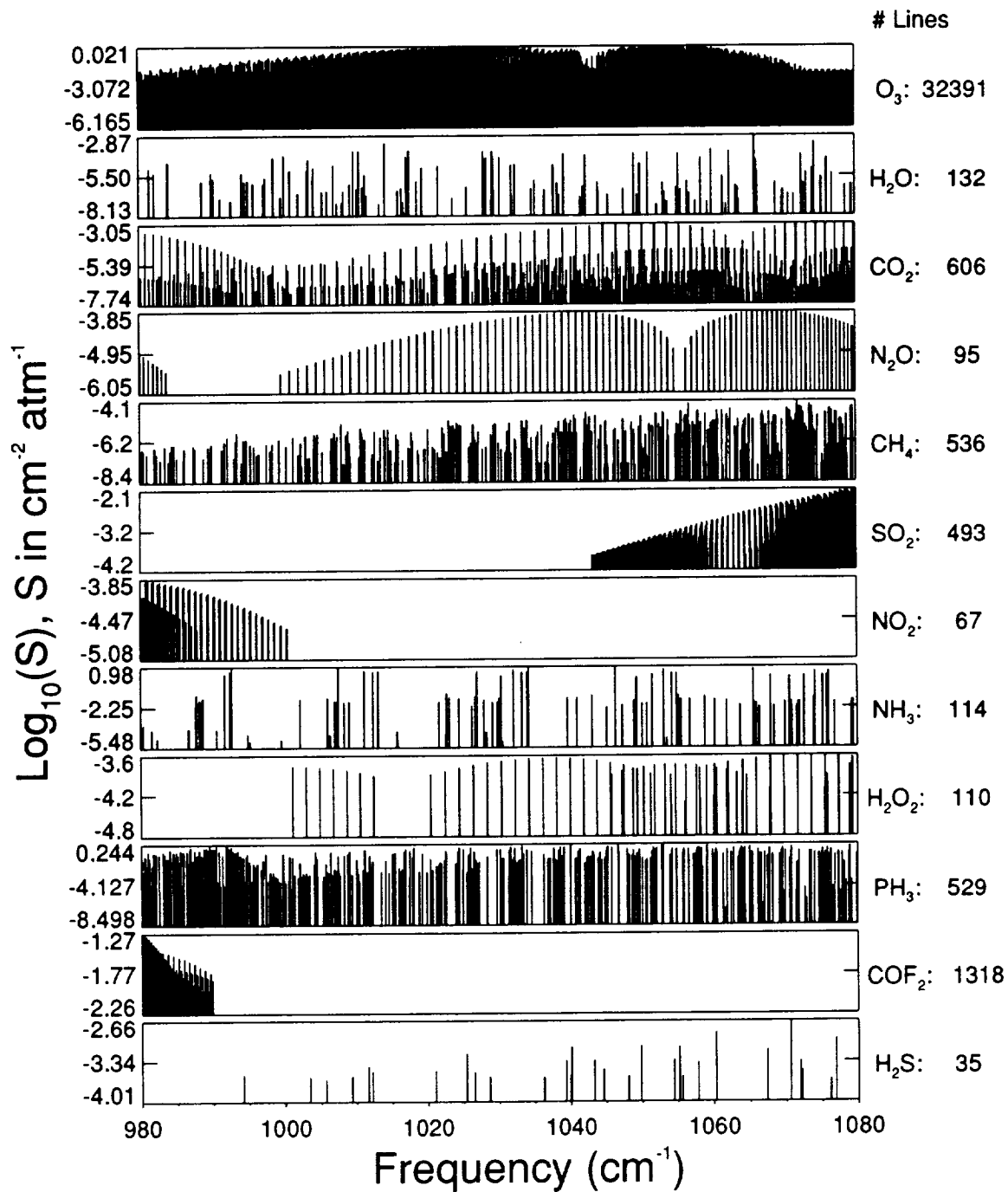


Figure 4.2: Spectral lines located within the 9.6 μm ozone band according to the 1992 AFGL compilation.

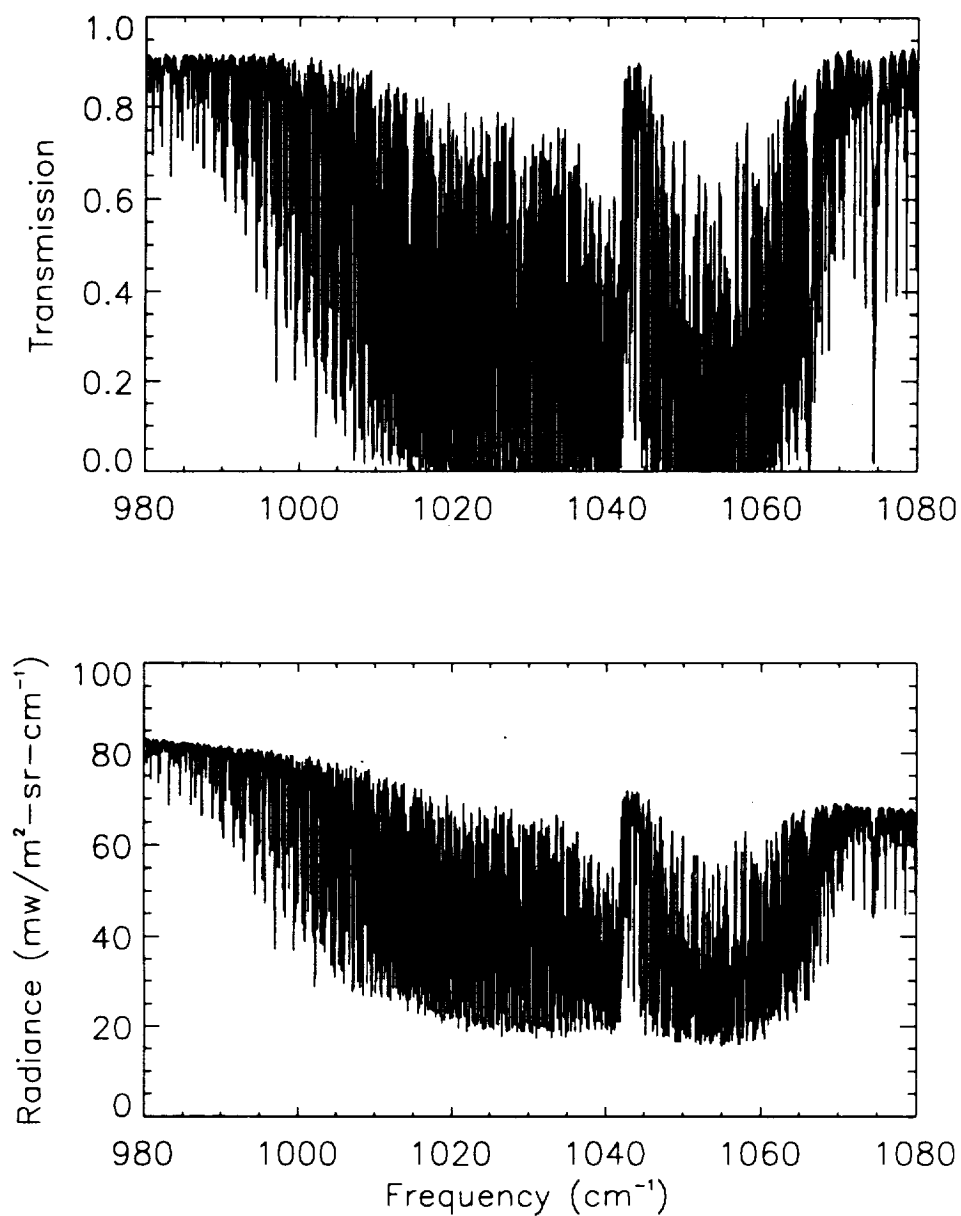


Figure 4.3: Atmospheric transmission and upwelling radiance computed at $.01\text{ cm}^{-1}$ resolution for the 1976 Standard Atmosphere. Calculations include spectral lines for all species contained within 1992 AFGL line parameter compilation except COF_2 and H_2S .

mospheric Signature Code (FASCODE). The 1992 Air Force Geophysics Laboratory (AFGL) atmospheric absorption line parameters compilation (Rothman, 1993) was utilized; it is based upon earlier releases (Rothman et al., 1987) and contains improved O_3 parameters for the $10\ \mu\text{m}$ spectral region (Flaud et al., 1990). Absorption lines in the $10\ \mu\text{m}$ spectral region are shown in Figure 4.2. Considering both line strengths and atmospheric concentrations, the most radiatively active molecules in this spectral region consist of: O_3 , H_2O , CO_2 , N_2O , and CH_4 . Since vertical profiles of SO_2 , NO_2 , NH_3 , H_2O_2 , and PH_3 were available in the 1976 Standard Atmosphere database (Anderson et al., 1986) used for this study, these species were also included in all radiance calculations. However, profiles of COF_2 and H_2S were not readily available and these two species were not included in the calculations since they are relatively insignificant in this spectral region. In addition, the radiance computations assume the surface to radiate as a perfect blackbody with an emissivity $\epsilon_s = 1$. Subsequent calculations of transmittances and corresponding radiances for the 1976 U.S. Standard Atmosphere showed a quite favorable comparison between the programs developed for this research and the FASCODE2 program unit alone. Figure 4.3 shows the atmospheric transmission and total upwelling radiance (including contributions from surface and atmosphere) calculated for the 1976 U.S. Standard Atmosphere within the $10\ \mu\text{m}$ spectral region.

Through perturbing the vertical distribution of atmospheric ozone in regions about which information content is desired and calculating the corresponding change in upwelling radiance, one can determine spectral regions desirable for instrument band placement as well as the approximate instrument spectral resolution necessary to observe the calculated radiance residuals. To this end, radiance calculations were performed for two additional atmospheric scenarios: First a case illustrating sensi-

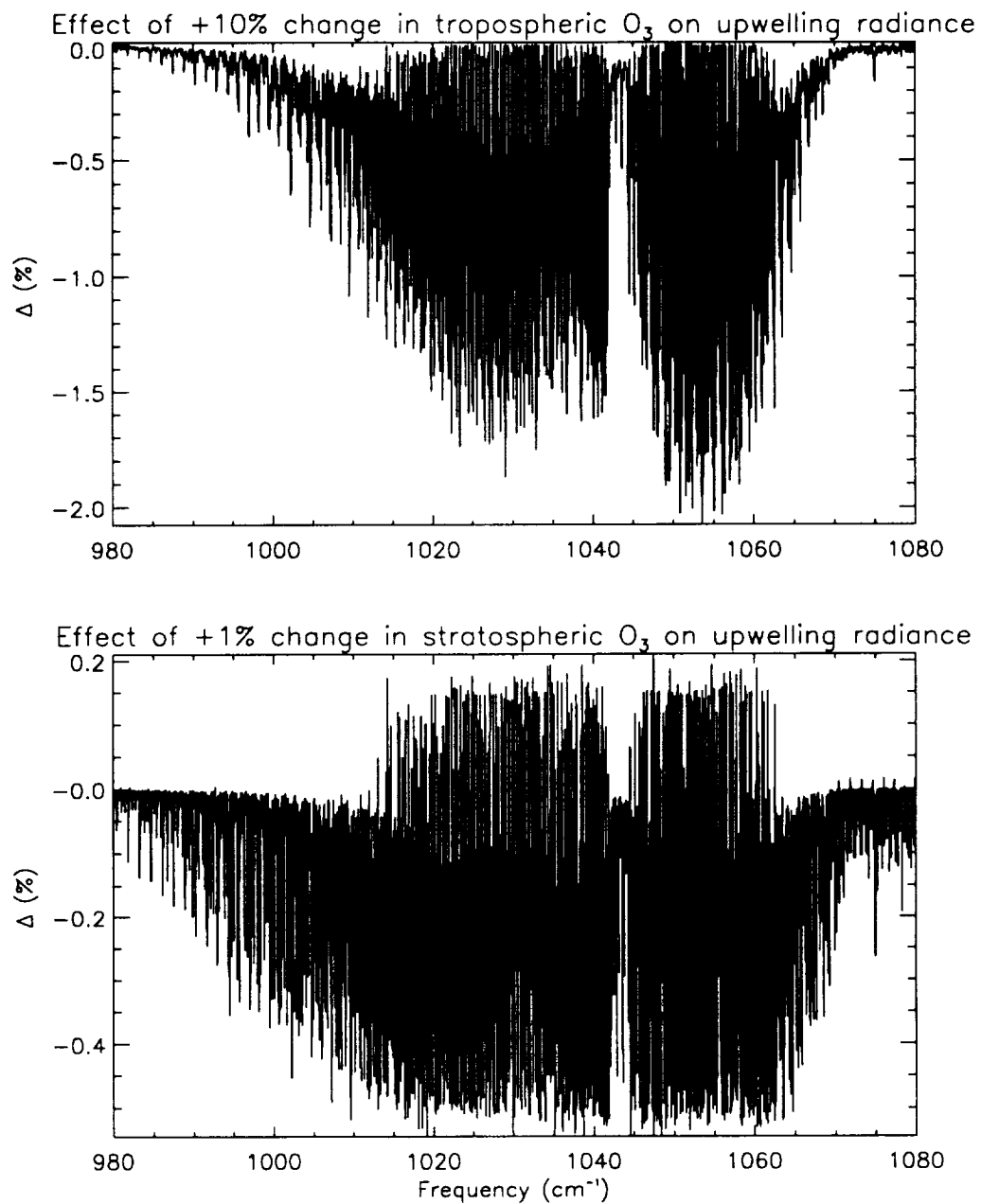


Figure 4.4: Effect of atmospheric ozone perturbations on upwelling radiance within the $9.6 \mu m$ band region.

tivity to a tropospheric ozone perturbation was simulated with a 10% increase in tropospheric ozone from the nominal (1976 U.S. Standard Atmosphere) vertical profile; analogously, calculations with a 1% increase in stratospheric ozone were performed to demonstrate sensitivity to stratospheric ozone changes. The ozone perturbation magnitudes were chosen to roughly correspond to a 1% change in total column ozone, an occurrence of which would be desirable to detect. Although the vertical uniformity and magnitude of these perturbations may deviate from reality, thus making the utility of the radiance residual magnitudes suspect, useful information pertaining to instrument spectral band placement and resolution can still be obtained. This simulation is summarized in Figure 4.4 with plots of the radiance residuals resulting from the above described ozone perturbations. As expected, maximum sensitivity to tropospheric changes can be found in the wings of the strongest ozone lines, for example, the $1048\text{-}1058\text{ cm}^{-1}$ spectral region. The negative radiance residuals result from the increased absorption reducing the large surface contribution to a greater extent than the radiance supplement from increased emission by atmospheric layers cooler than the surface temperature. The positive radiance residuals in the stratospheric perturbation case are due to the dominance of the change in atmospheric emission in spectral regions centered about strong lines which therefore minimize the surface contribution; since the atmosphere is essentially opaque in these spectral regions, tropospheric perturbations cannot be seen and thus yield radiance residuals ≈ 0 in these same spectral locations. Additionally, the sensitivity to changes in atmospheric ozone goes to zero as one approaches the wings of the band where the absorption lines are less intense. From Figure 4.4, one can conclude that somewhere in the $1048\text{-}1058\text{ cm}^{-1}$ spectral region would be ideally suited for placement of the instrument band (or channels) to maximize information content for both the

tropospheric and stratospheric ozone measurements.

The character of the radiance residual plots can be better appreciated by expanding a small spectral region along with the corresponding ozone spectral lines. This is shown in Figure 4.5 for a 2 cm^{-1} interval near the center of the $1048\text{-}1058\text{ cm}^{-1}$ spectral region of interest; transmission through the nominal standard atmosphere is also illustrated for a reference of atmospheric opacity as well as an indicator of relative importance of the surface contribution to the upwelling radiance. As Figure 4.5 demonstrates, regions in the wings of strong absorption lines have the largest sensitivity to changes in tropospheric ozone and would thus be the best spectral locations for placement of tropospheric channels; for example, 1051.8 cm^{-1} , or 1052.1 cm^{-1} for a higher tropospheric sensitivity along with a higher stratospheric sensitivity contribution. These plots show that an instrument resolution $\sim .07\text{ cm}^{-1}$ should be sufficiently narrow for the tropospheric ozone measurement in addition to enabling determination of stratospheric ozone content.

4.4 Selection of Double-Etalon Instrument Characteristics

As discussed in appendix A, one is desirous of obtaining the best quality etalon plates possible to maximize the overall defect finesse N_D . Many manufacturers (i.e., Imperial College Optical Systems (ICOS), LTD) are currently able to produce large etalons with working apertures up to 150 mm in diameter of very high plate quality with $N_D \approx 50$ for the IR spectral region. However, the overall finesse may be severely degraded during the mounting process in preparation for space flight (Hays, private communication, 1993). An overall defect finesse of 25% below the nominal value given above, or $N_D \approx 37.5$, will be arbitrarily assumed in this research. This author feels that if further degradation of N_D is anticipated in preparation for space

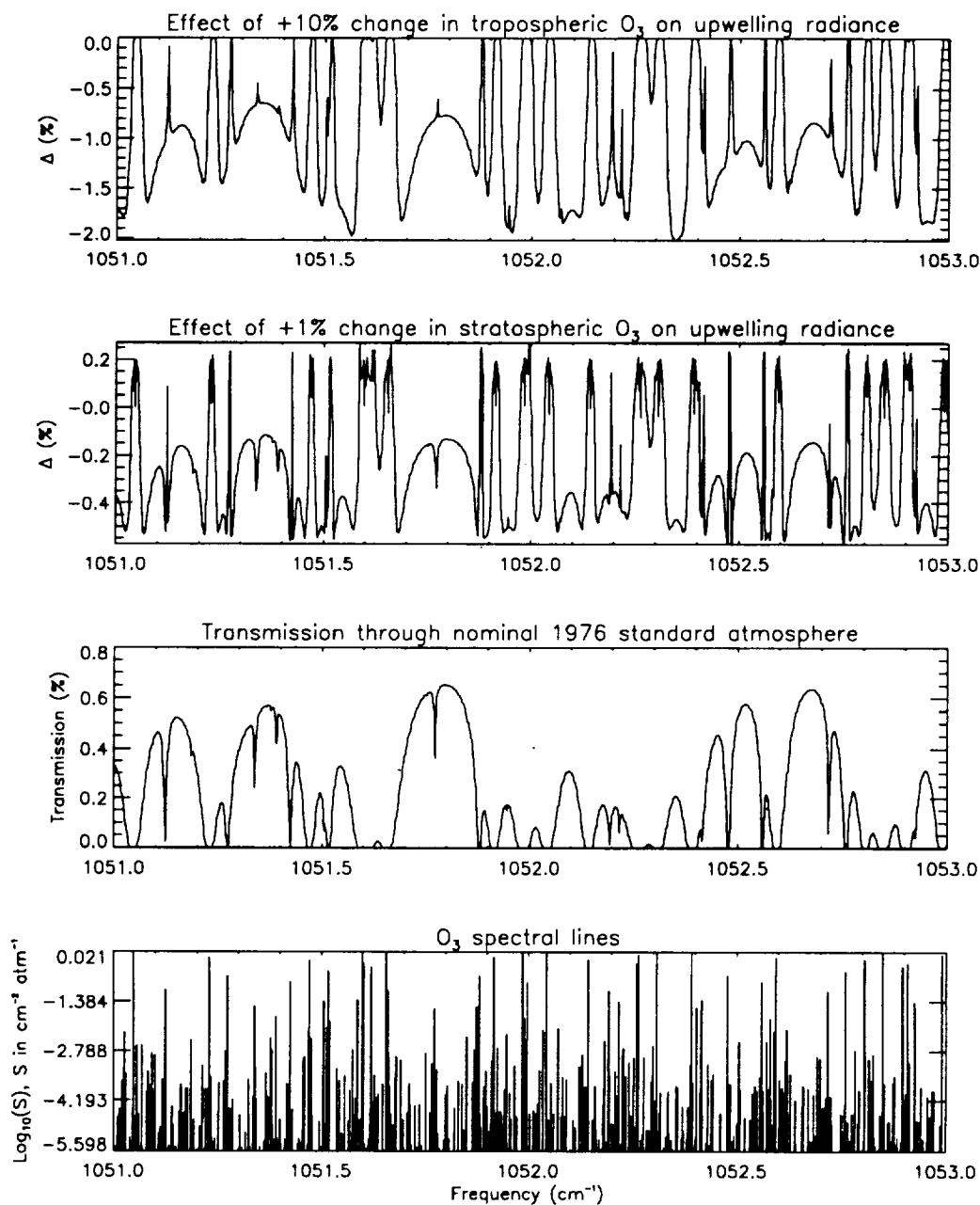


Figure 4.5: Effect of atmospheric ozone perturbations on upwelling radiance in a narrow spectral region of strong ozone lines. The transmission through an unperturbed atmosphere and the ozone spectral line distribution are also illustrated.

flight, then an alternative optical component integration procedure should be devised and/or the application of solid etalons (in place of the presently assumed fixed-gap etalons) might be of consideration for this particular research application. Due to lack of apriori knowledge regarding the finesses corresponding to individual defects, it will be assumed that random surface imperfections, spherical curvature, and lack of parallelism all contribute to the total defect of the plates with equal defect finesses. Using equation A.18, the individual defect finesses can be calculated to be $N_{D_G} = N_{D_S} = N_{D_P} \approx 65.0$. In addition, the defect distribution functions (see, for example, Section A.3) will be assumed to be symmetrical about their means and to have mean values of zero ($\mu_D = 0$). To maximize throughput of a single etalon, the reflectivity finesse should be chosen such that $N_R \leq N_D$, which (for $N_D \approx 37.5$) implies plate reflectivities $R \leq .92$. However, the peak transmittance in a multiple etalon system is lower than that of a single-etalon instrument, unless the plate reflectivities are lower in the multi-etalon configuration. In addition, the multi-etalon transmittance decreases sharply and becomes more sensitive to plate defects for high plate reflectivities. Roesler (1969) discussed the effect of plate defects on the transmission function for a poly-etalon FPI. For a double-etalon series configuration, he suggested a reflectivity finesse equal to the defect finesse lowered by a factor of .644 (i.e., $N_R = .644N_D$) in order for the system transmittance to not be degraded from the multi-etalon configuration; for the design under consideration, this implies plate reflectivities $R \approx .88$. However, the use of lower plate reflectivities can decrease the spectral resolution and increase the amount of parasitic or off-band light transmitted.

As mentioned in section 4.3, the desired spectral resolution is $\delta_{F_E} \approx .07 \text{ cm}^{-1}$. Since the double-etalon spectral resolution is determined primarily by the high-resolution etalon (HRE), we can use equation A.24 to estimate the necessary free

spectral range for the HRE; the effective finesse can be computed from equation A.25 by assuming the aperture finesse to be equal to the reflectivity finesse (i.e., $N_L \approx N_R$). If we consider, for example, the coating reflectivities for both etalons to be $R = .90$ ($N_R \approx 29.8$), this would imply an effective finesse $N_E \approx 18.4$ and a free spectral range of $\nu_{fsr} \approx 1.29$ for the HRE. The ν_{fsr} of the low-resolution etalon (LRE) must be chosen such that single-order passband transmission is achieved in addition to minimizing the amount of unwanted energy through the optical system. This unwanted energy is frequently referred to as white or parasitic light and can be defined by the ratio of all energy transmitted through the system outside the HRE central order to the energy transmitted within the HRE central order (Skinner et al., 1987). The white light W can then be expressed mathematically as

$$W = \frac{\int_0^\infty T_F(\nu)T_D(\nu)d\nu - \int_{\nu_0 - \frac{HRE\nu_{fsr}}{2}}^{\nu_0 + \frac{HRE\nu_{fsr}}{2}} T_F(\nu)T_D(\nu)d\nu}{\int_{\nu_0 - \frac{HRE\nu_{fsr}}{2}}^{\nu_0 + \frac{HRE\nu_{fsr}}{2}} T_F(\nu)T_D(\nu)d\nu}, \quad (4.17)$$

where $T_F(\nu)$ is the bandpass filter transmission function, and $T_D(\nu)$ is the double-etalon instrument transfer function which includes broadening effects and reflections between etalons (equation 3.4). As illustrated in Figure 3.3, placement of a loss medium in between the etalons can significantly reduce the undesirable increased white light effect from inter-etalon reflections at the cost of slightly decreasing peak transmission; accordingly, a loss medium transmission value of $\tau_m = .95$ will be assumed in this research. To optimize the instrument's spectral isolation capability and thus minimize the amount of white light, it is desired that the bandpass filter be as narrow as technologically feasible; this also implies a reduction in peak transmission. For this spectral region, ultra-narrow bandpass filters with a full-width at half-maximum on the order of .4% of the central wavenumber ($\text{FWHM} \sim .4\%\nu_0$) have been produced (i.e., by the Optical Coatings Laboratory, Inc., (OCLI) for the

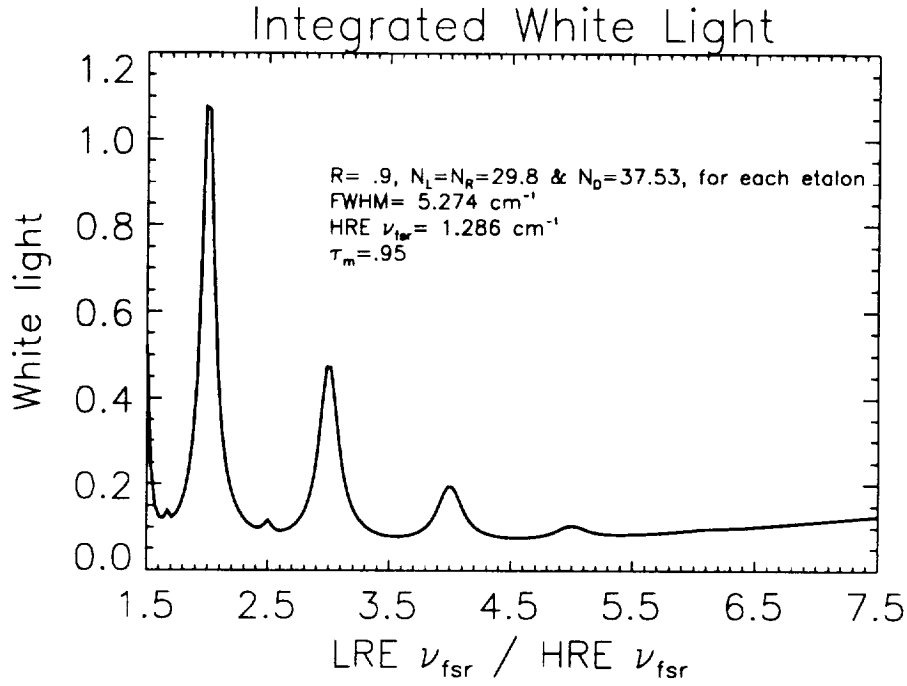


Figure 4.6: White light transmitted through a double-etalon Fabry-Perot system. An ultra-narrow bandpass filter of Gaussian shape is included. A medium with transmission equal to .95 is assumed to be in between the etalons to reduce the effect of incoherent reflections.

CLAES instrument aboard the UARS platform); the corresponding peak transmission is $T_{F_{max}} \approx 65\%$. An ultra-narrow bandpass filter with $\text{FWHM} = .5\% \nu_0$ and $T_{F_{max}} = 65\%$ will be assumed in this study; such a filter can be modeled by a Gaussian shape. Thus the filter transmission will be represented by

$$T_F(\nu) = \exp \left[\frac{-(\nu - \nu_0)^2}{2\sigma^2} \right], \quad (4.18)$$

where the standard deviation σ of a Gaussian can be given by $\sigma = \frac{\text{FWHM}}{[2\sqrt{2\ln 2}]}$.

Evaluating equation 4.17 using the instrument specifications noted above and a range of LRE free spectral range values yields white light amounts depicted in Figure 4.6. Although near-vernier gap ratios enable the highest spectral resolution, they also correspond to high white light transmission since the HRE parasitic peaks

R	$T_{D_{max}}$	W_{min}	$\delta_{FE}(cm^{-1})$	$\nu_{fsr_1}(cm^{-1})$	$\nu_{fsr_2}(cm^{-1})$
0.850	0.783	0.329	0.067	0.898	3.223
0.860	0.758	0.260	0.068	0.959	3.441
0.870	0.735	0.201	0.068	1.027	3.685
0.880	0.702	0.152	0.068	1.103	3.958
0.890	0.668	0.111	0.068	1.189	4.267
0.900	0.634	0.080	0.066	1.286	4.615
0.910	0.580	0.055	0.068	1.396	5.009
0.920	0.536	0.037	0.068	1.520	5.455
0.930	0.471	0.024	0.067	1.660	5.957
0.940	0.405	0.015	0.066	1.816	6.515
0.950	0.329	0.010	0.066	1.985	7.123

Table 4.2: Simulation demonstrating effect of varying plate reflectivity while maintaining fixed spectral resolution on peak double etalon transmission $T_{D_{max}}$ and minimum white light W_{min} through the system.

Instrument Design Specifications		
Ultra Narrow Bandpass Filter:		
Gaussian shape assumed; central wavenumber, $\nu_0 = 1054.73 \text{ cm}^{-1}$;		
FWHM = .5 % $\nu_0 = 5.274 \text{ cm}^{-1}$; maximum transmittance, $T_{F_{max}} = 65\%$		
Etalons:		
	HRE (etalon #1)	LRE (etalon #2)
gap t	0.3289 cm	0.09157 cm
ν_{fsr}	1.52 cm^{-1}	5.46 cm^{-1}
N_D	37.53	37.53
reflectivity R	0.92	0.92
N_R	37.66	37.66
$\tau_{etalon_{max}}$	0.732	0.732
Additional parameters:		
Aperture finesse $N_L = 37.66$		
Transmission through medium between etalons $\tau_m = .95$		
Transmission through remaining optical elements $\tau_o = .85$		
Overall effective system finesse $N_E \approx 22.35$		
Effective resolution $\delta_{F_E} = 0.068 \text{ cm}^{-1}$		
White light W = .037		
Instrument peak transmittance $\tau_i = 0.28$		
Etalon plates diameter $D_e \approx 13.3 \text{ cm}$		
Telescope diameter $d_T = 9.0 \text{ inches}$		
Nadir footprint size $d_{NF} = 50.0 \text{ km}$		

Table 4.3: Double-etalon instrument design specifications.

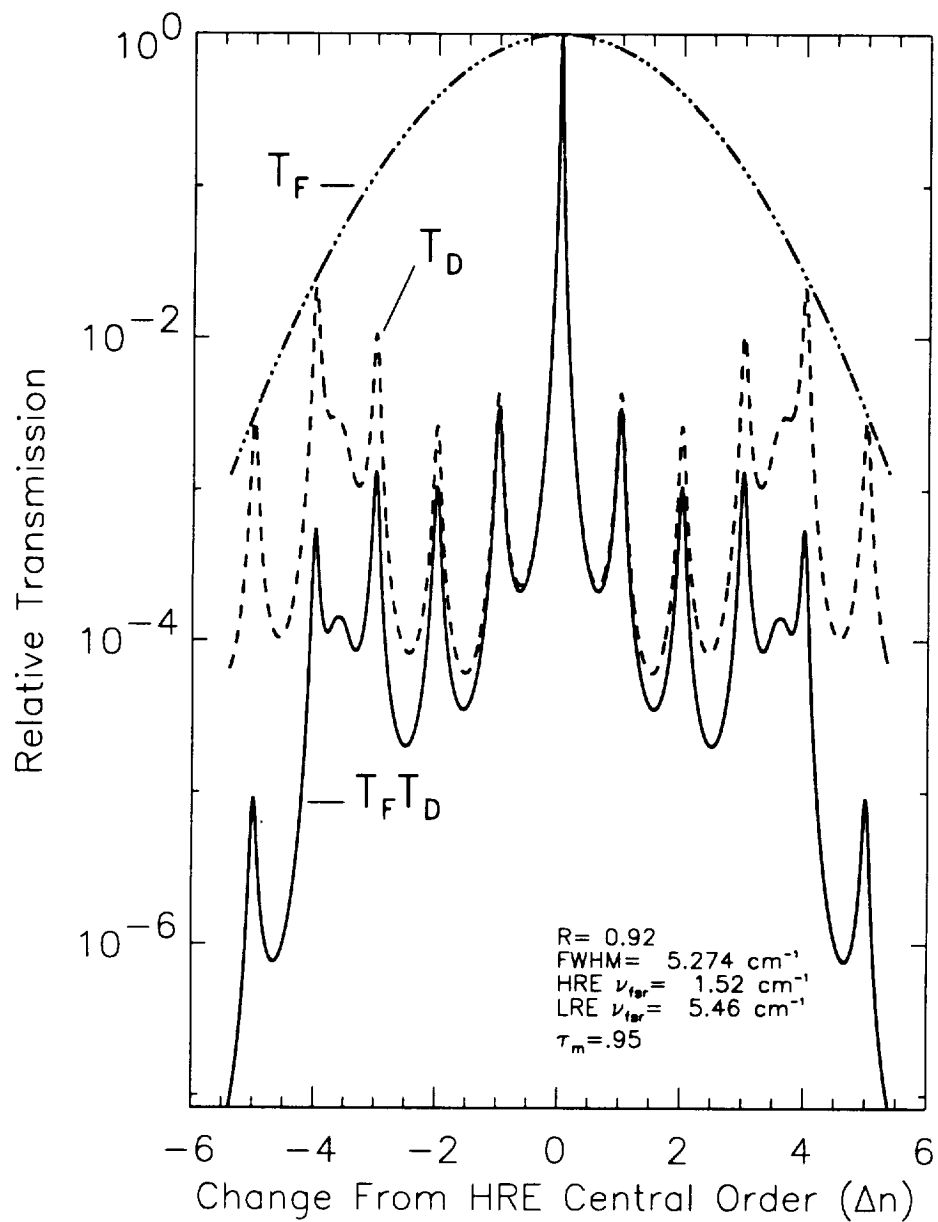


Figure 4.7: Illustration of ultra-narrow bandpass filter function, broadened double-etalon transmission function, and overall instrumental transfer function of the central wavenumber for an FPI with the chosen specifications.

are not blocked. The effect of the bandpass filter can be seen in Figure 4.6 as the local maxima decrease in intensity for increasing integer spacer ratios. In this simulation, the white light minimum ($W_{min} \approx .08$) occurs for a free spectral range ratio $\frac{\nu_{fsr2}}{\nu_{fsr1}} \approx 3.59$. A design compromise exists between peak throughput and white light; since white light can be reduced significantly at the expense of a relatively small drop in throughput, if the plate reflectivities are increased while maintaining the overall resolution approximately constant (and thus allowing the free spectral ranges to get larger). This is demonstrated in Table 4.2. Based on these results, a plate reflectivity of $R = .92$ ($N_R \approx 37.7$), with free spectral ranges $\nu_{fsr1} = 1.52 \text{ cm}^{-1}$ and $\nu_{fsr2} = 5.46 \text{ cm}^{-1}$ will be implemented into this design and subsequent instrument simulations. Table 4.3 summarizes the instrument design specifications discussed above, with an additional loss factor of $\tau_o = .85$ incorporated to account for decrease in transmission from the remaining optical elements not yet considered (i.e., beamsplitters, lenses, etc.). Note that the choice of central wavenumber ν_0 will be justified in the next section, while diameters of the etalon plates, entrance telescope, and nadir footprint will be arrived upon in section 4.6. The spectral character of the resulting instrument function is illustrated in Figure 4.7; the product of the bandpass filter and double-etalon transmission functions ($T_F T_D$) shows the effective central wavenumber instrument transfer function.

4.5 Band Placement and Choice of Spectral Channels

The $1048\text{-}1058 \text{ cm}^{-1}$ spectral region was selected in section 4.3 as having potential for placement of the instrument measurement band in order to enable the desired atmospheric ozone observation. Spectral locations for the individual channels will be determined in this section based upon simulations performed using the

instrument function designed in section 4.4. However, some of the signal related equations introduced in section 2.5 must first be re-expressed to reflect the nadir-viewing geometry and the Fabry-Perot instrument function formulation. Accordingly, the measured signal $S(\nu)$ (equation 2.9), the measured signal normalized to instrument transmission $S_N(\nu)$ (equation 2.12), and the energy contribution functions $C(\nu, p)$ (equation 2.11), can be expressed as

$$S(\nu) = \int_{\nu_{low}}^{\nu_{high}} T_D(\nu - \nu') T_F(\nu') I(\nu') d\nu', \quad (4.19)$$

$$S_N(\nu) = \frac{\int_{\nu_{low}}^{\nu_{high}} T_D(\nu - \nu') T_F(\nu') I(\nu') d\nu'}{\int_{\nu_{low}}^{\nu_{high}} T_D(\nu - \nu') T_F(\nu') d\nu'}, \text{ and} \quad (4.20)$$

$$C(\nu, p) = \frac{\int_{\nu_{low}}^{\nu_{high}} T_D(\nu - \nu') T_F(\nu') B(\nu', T(p)) \frac{\partial \tau(\nu', p)}{\partial \ln p} d\nu'}{\int_{\nu_{low}}^{\nu_{high}} T_D(\nu - \nu') T_F(\nu') d\nu'}, \quad (4.21)$$

respectively; the ν_{low} and ν_{high} represent the low and high wavenumber cutoffs, respectively, corresponding to the bandpass filter transmission. To ensure the existence of necessary information content, ozone energy contributions functions (O_3 ECFs) with their wavenumber and pressure dependence must be closely examined. The spectral lines of interfering species and their corresponding effect on the O_3 ECFs must also be considered. In addition, the signal sensitivity to perturbations in atmospheric ozone amount needs to be incorporated into the band and channel selection process.

Figure 4.8 shows all spectral lines located within the 1048-1058 cm^{-1} region. For this same spectral region, the O_3 ECFs and how they are affected from including interfering species are illustrated in Figure 4.9; the effect on the O_3 ECFs from including interfering species can be calculated from

$$\Delta[C(\nu, p)|_{O_3}] = \frac{(C(\nu, p)|_{all\ species} - C(\nu, p)|_{O_3\ only})}{C(\nu, p)|_{O_3\ only}}. \quad (4.22)$$

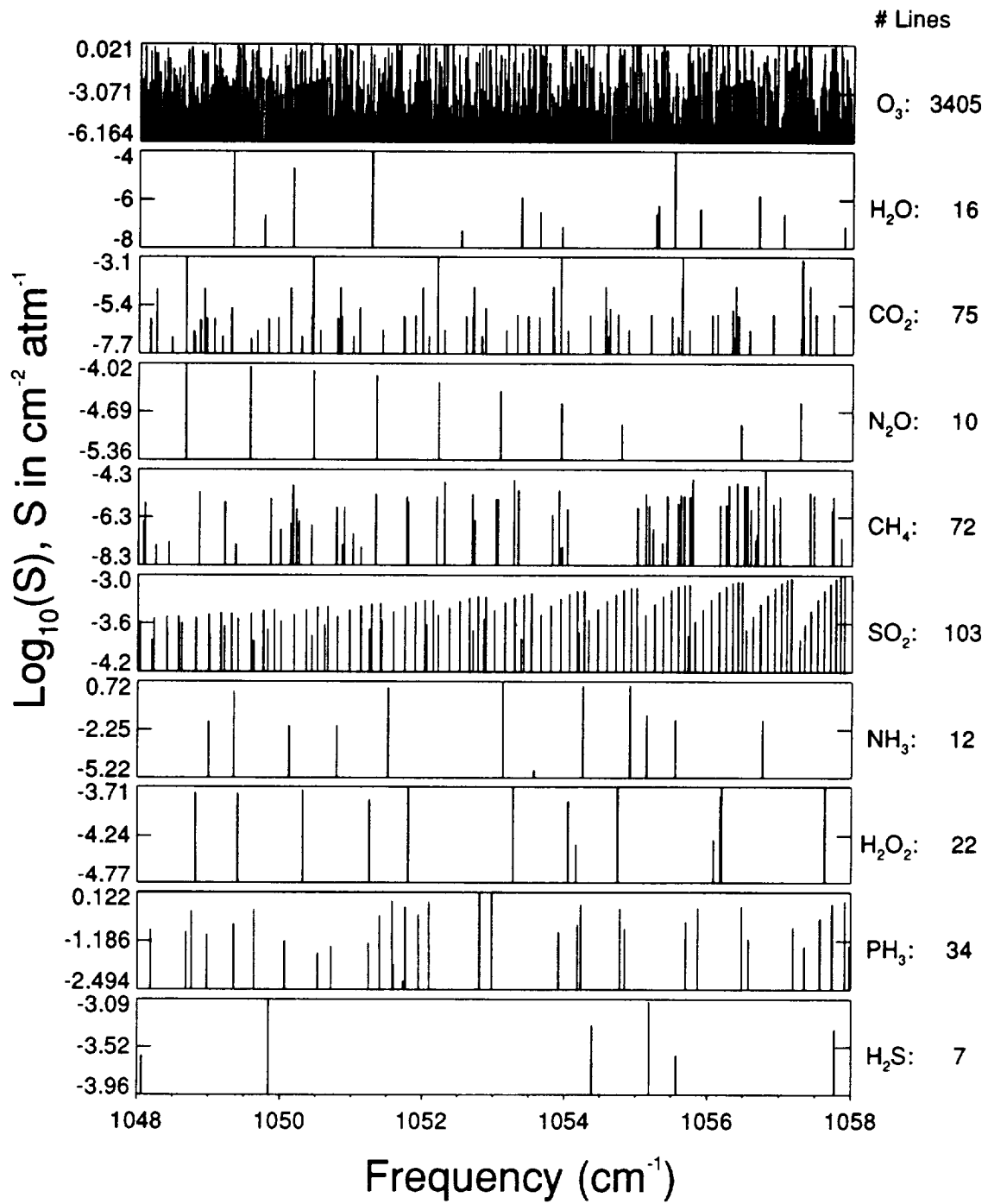


Figure 4.8: Spectral lines located within a portion of the 9.6 μm ozone band according to the 1992 AFGL compilation.

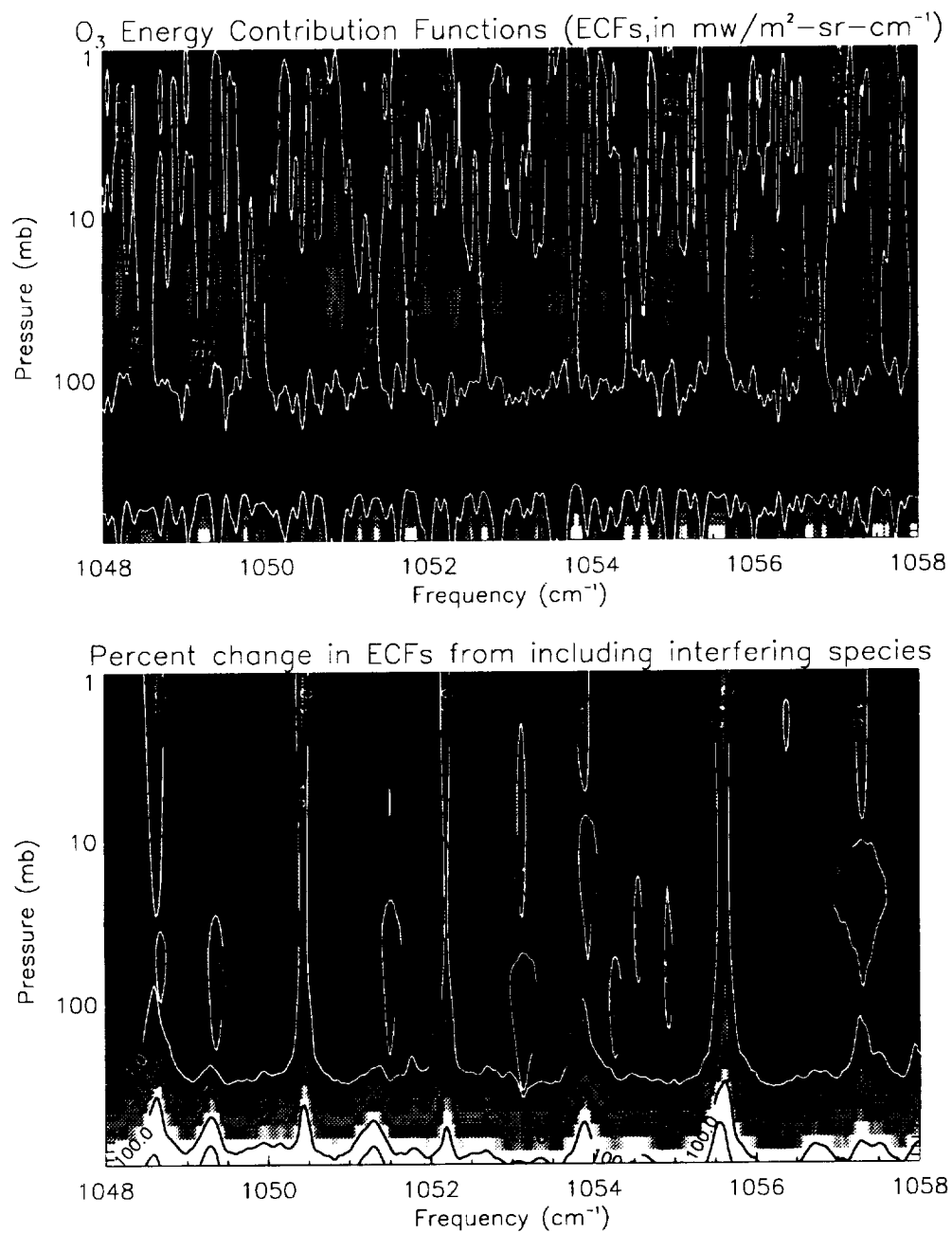


Figure 4.9: O₃ energy contribution functions and how they are affected from including interfering species for the 1048-1058 cm^{-1} region.

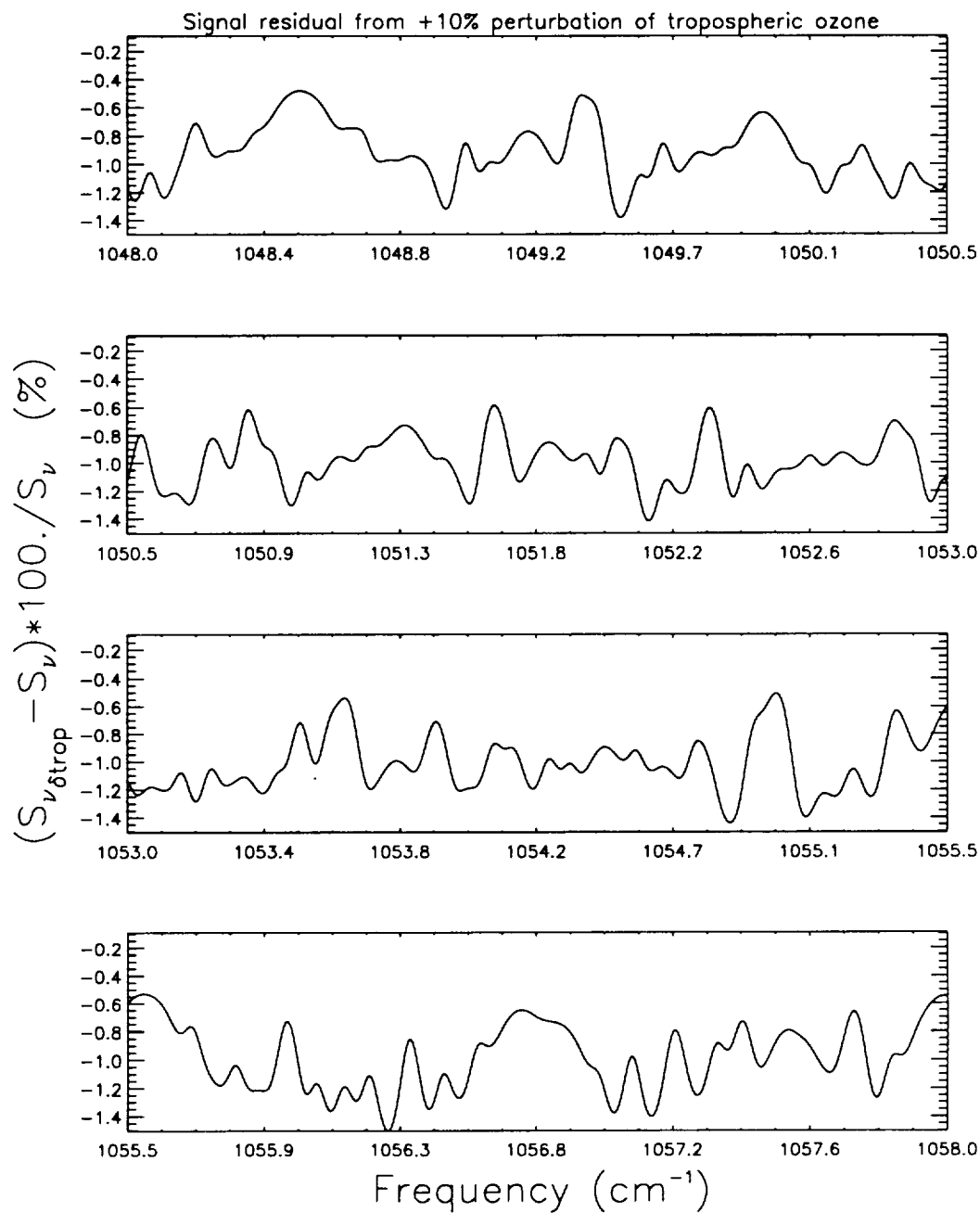


Figure 4.10: Signal sensitivity to a 10 percent increase in tropospheric ozone.

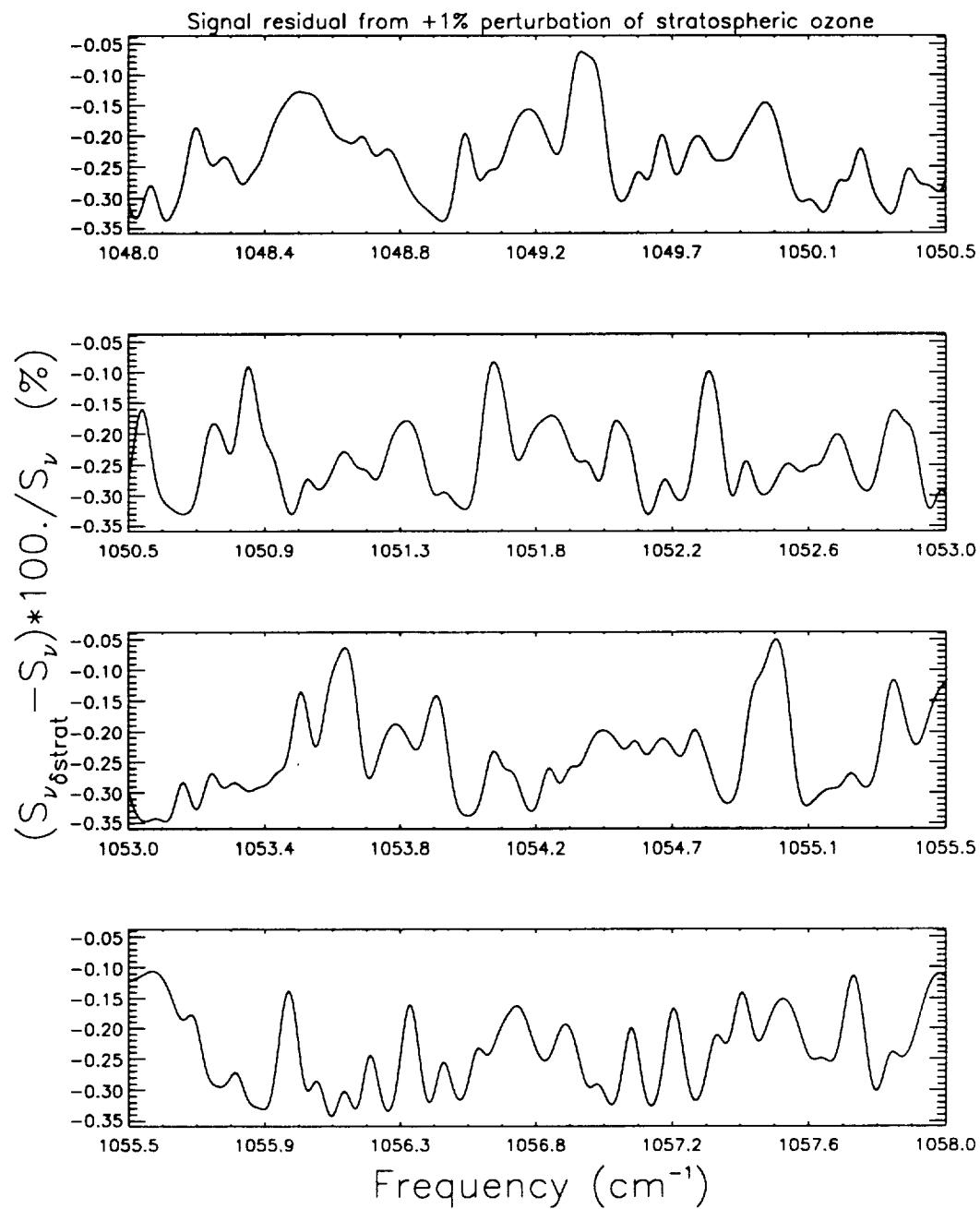


Figure 4.11: Signal sensitivity to a 1 percent increase in stratospheric ozone.

The O_3 ECFs in Figure 4.9 show that spectral regions can be chosen to maximize stratospheric or tropospheric information content in the measured signal; the effect of interfering species' spectral lines and the water vapor continuum can easily be observed in the corresponding $\Delta[C(\nu, p)|_{O_3}]$ contour/image color display product. Applying equation 4.20 to the upwelling radiance data sets generated during the atmospheric ozone perturbation simulations discussed in section 4.3 (see, for example, Figure 4.4), one can calculate the corresponding instrument signal sensitivity to changes in atmospheric ozone content. For the 1048-1058 cm^{-1} spectral region being considered, the signal sensitivities to a 10% increase in tropospheric ozone and to a 1% increase in stratospheric ozone are illustrated in Figures 4.10 and 4.11, respectively. As the plots demonstrate, convolution with the instrument function serves to smooth out the radiance response to the ozone perturbations. It is also interesting to note that peak sensitivities to the different perturbations occur spectrally near each other, due to the mutual close proximity to strong absorption lines; this implies that both the tropospheric and stratospheric measurements can probably be achieved with the instrument band covering a narrow spectral region.

The information contained within Figures 4.9, 4.10, and 4.11 was scanned thoroughly to determine regions suitable for instrument band placement. Decision criteria were established to identify spectral regions possessing high sensitivity to ozone variation (especially in the troposphere), while minimizing the impact of interfering species and maximizing the vertical extent of the O_3 information content. In addition, based on instrument design considerations, it was desired to find these features in one narrow spectral region approximately 1 cm^{-1} wide. Although several regions had potential, the 1054.23-1055.23 cm^{-1} spectral region meets the desired instrument band criteria and was determined to be the most suitable from the can-

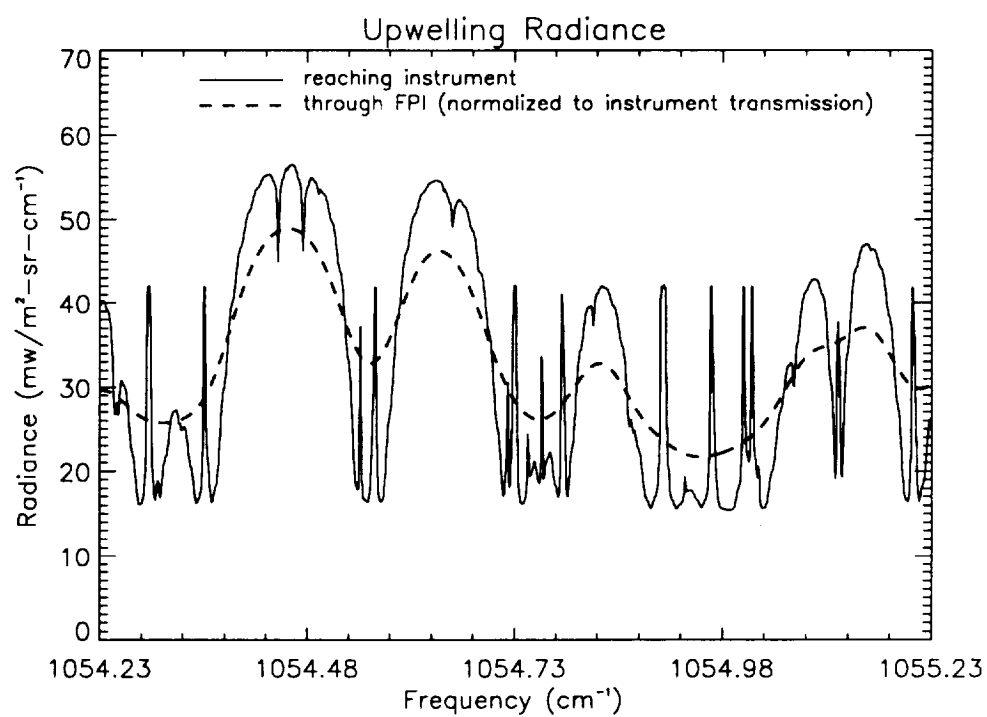


Figure 4.12: Upwelling radiance reaching and received by the instrument within the selected band region.

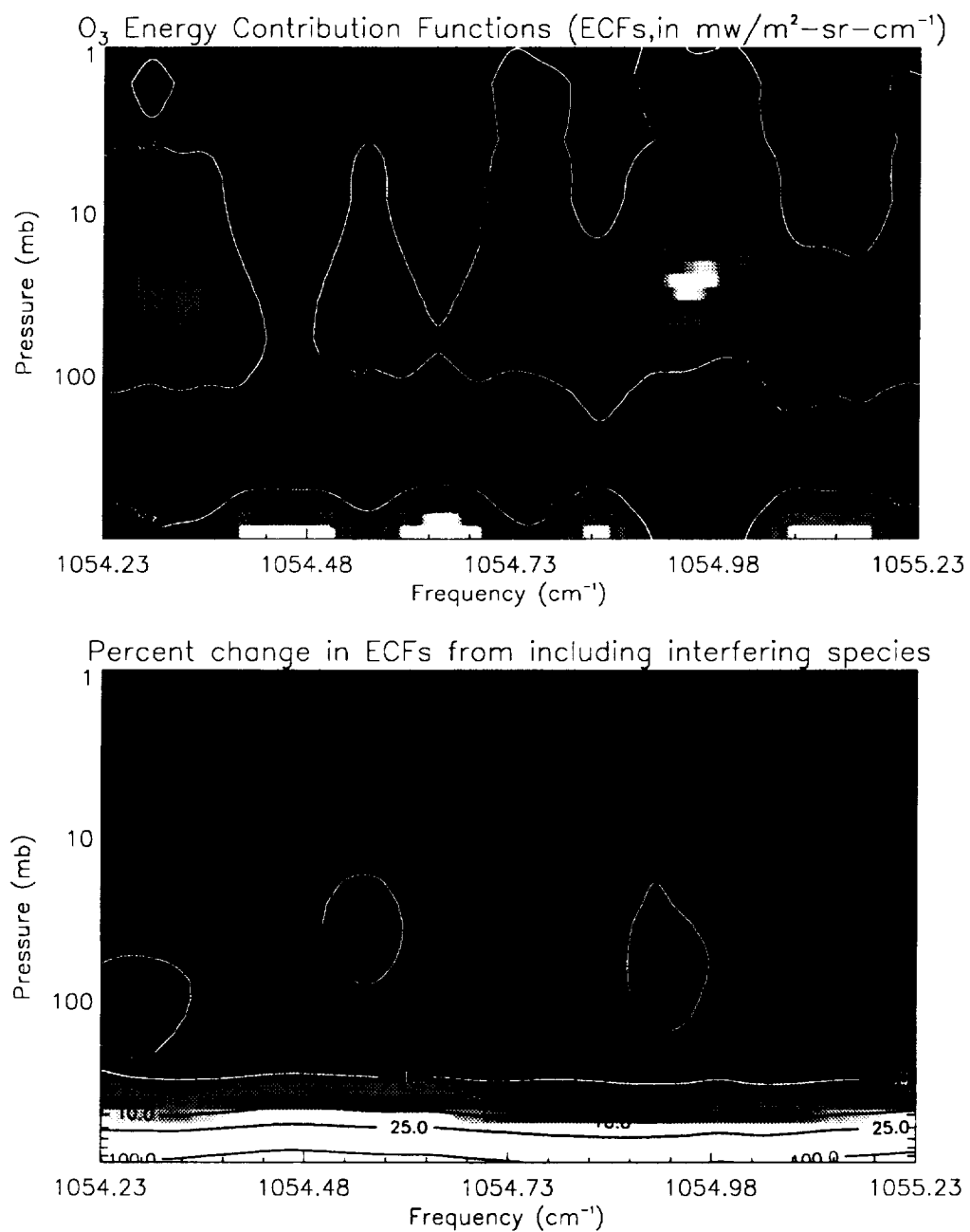


Figure 4.13: O_3 energy contribution functions and how they are affected from including interfering species for the instrument band spectral region.

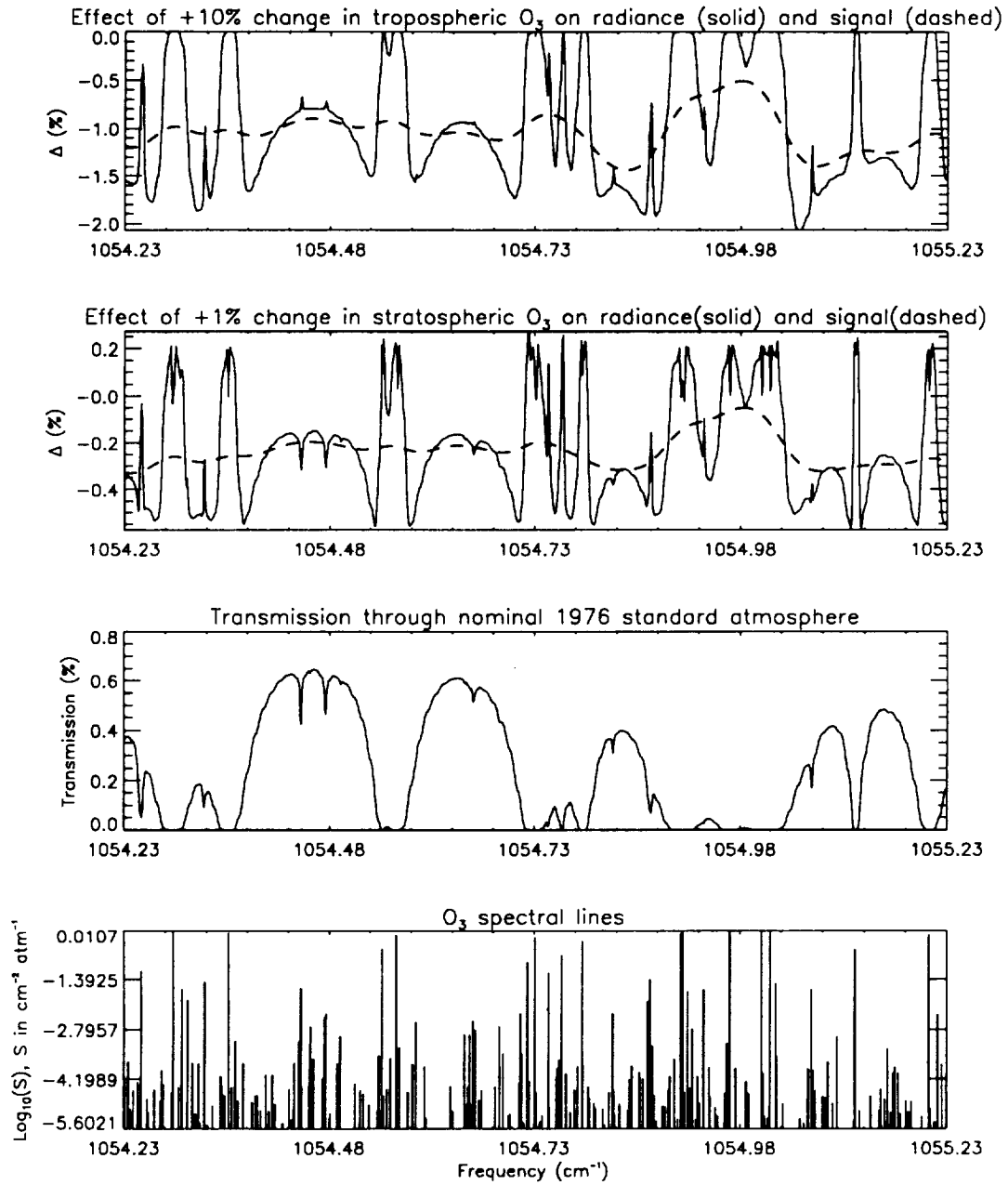


Figure 4.14: Effect of atmospheric ozone perturbations on the radiance and signal within the instrument band spectral region. The transmission through an unperturbed atmosphere and the ozone spectral line distribution are also illustrated.

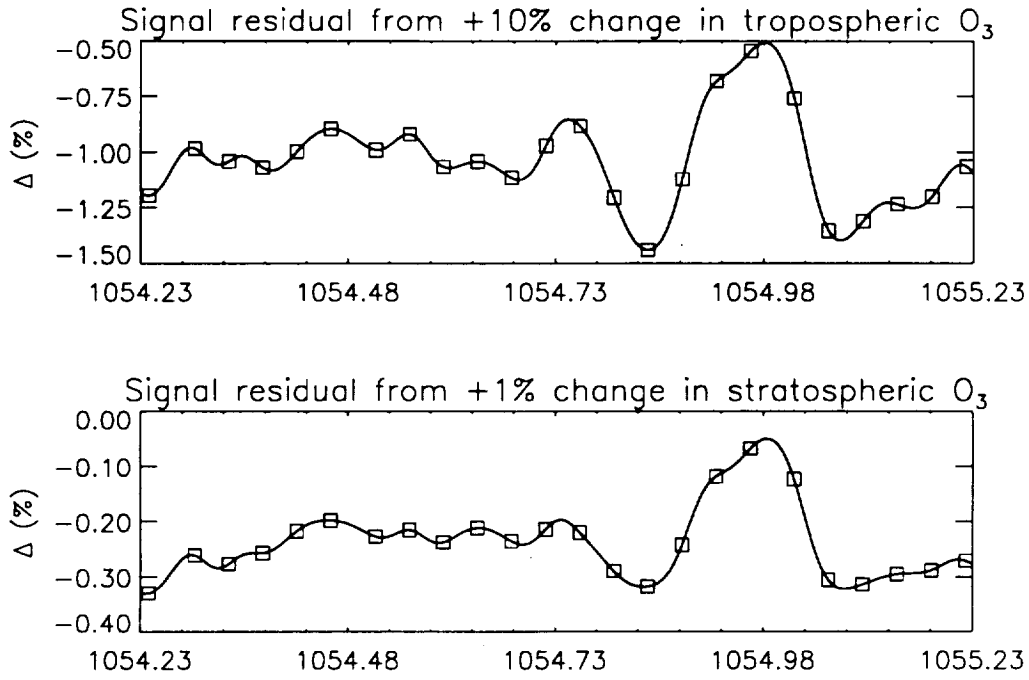


Figure 4.15: Signal sensitivities to a 10 percent increase in tropospheric ozone and to a 1 percent increase in stratospheric ozone are shown with the selected instrument channels.

didate regions. Figure 4.12 shows the upwelling radiance in the selected band region; the resultant smoothing effect of the instrument function (due to the measurement process) is also illustrated. Analogous to Figure 4.9, Figure 4.13 depicts the O_3 ECFs and the corresponding effect from including interfering species within the instrument band spectral region. Additionally, Figure 4.14 is similar to Figure 4.5, however, it represents the instrument band region and includes the signal sensitivity to O_3 changes superimposed over that for the upwelling radiance.

The next step is to choose channel locations. Since the primary interferant effect within the band region is due to the spectrally near-uniform H_2O vapor continuum (see Figure 4.13), there are no portions that must be avoided and it is desirable to sense the whole region. The channels are defined by the spatial distribution of detec-

Normalized O_3 energy contribution functions

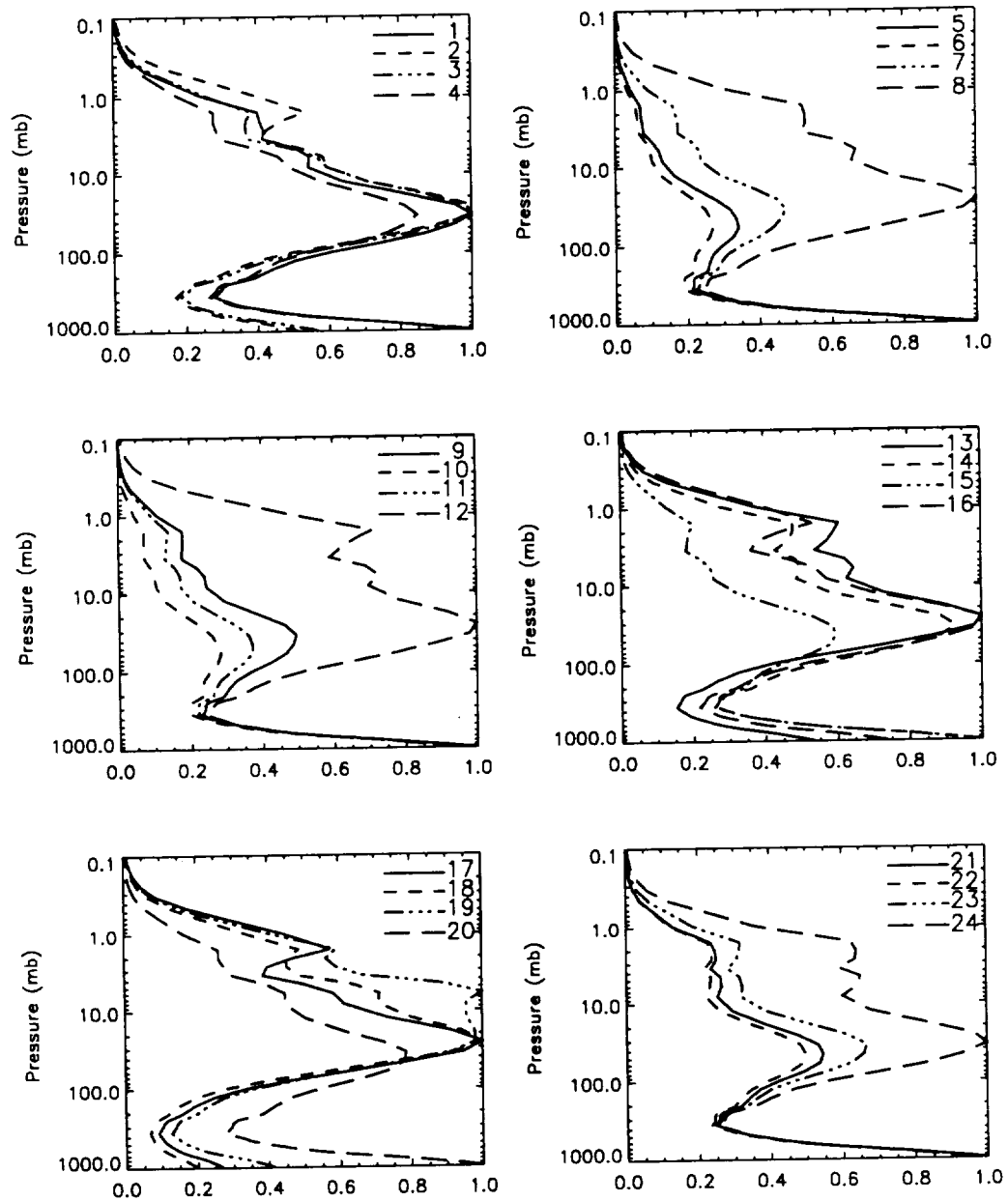


Figure 4.16: Energy Contribution Functions (ECFs) for the selected channels normalized to their peak values for an ozone only atmosphere.

channel	ν_{0_c} (cm^{-1})	$\bar{p}(mb)$ of peak layer	$[O_3ECF]_{max}$ ($mw\ m^{-2}sr^{-1}/cm^{-1}$)
1	1054.240	33.951	3.523
2	1054.295	33.951	4.049
3	1054.336	33.951	3.985
4	1054.377	983.597	3.634
5	1054.418	983.597	5.703
6	1054.459	983.597	6.058
7	1054.513	983.597	4.845
8	1054.554	25.416	3.335
9	1054.595	983.597	4.908
10	1054.636	983.597	6.465
11	1054.677	983.597	5.812
12	1054.718	25.416	3.322
13	1054.759	25.416	4.134
14	1054.800	983.597	3.767
15	1054.841	983.597	5.135
16	1054.882	25.416	3.925
17	1054.923	25.416	4.880
18	1054.964	25.416	5.085
19	1055.016	5.959	4.116
20	1055.057	983.597	4.146
21	1055.098	983.597	5.119
22	1055.139	983.597	5.265
23	1055.180	983.597	4.432
24	1055.221	33.951	3.400

Table 4.4: Characteristics of the O_3 ECFs for the instrument channels.

ECF change from including interfering species

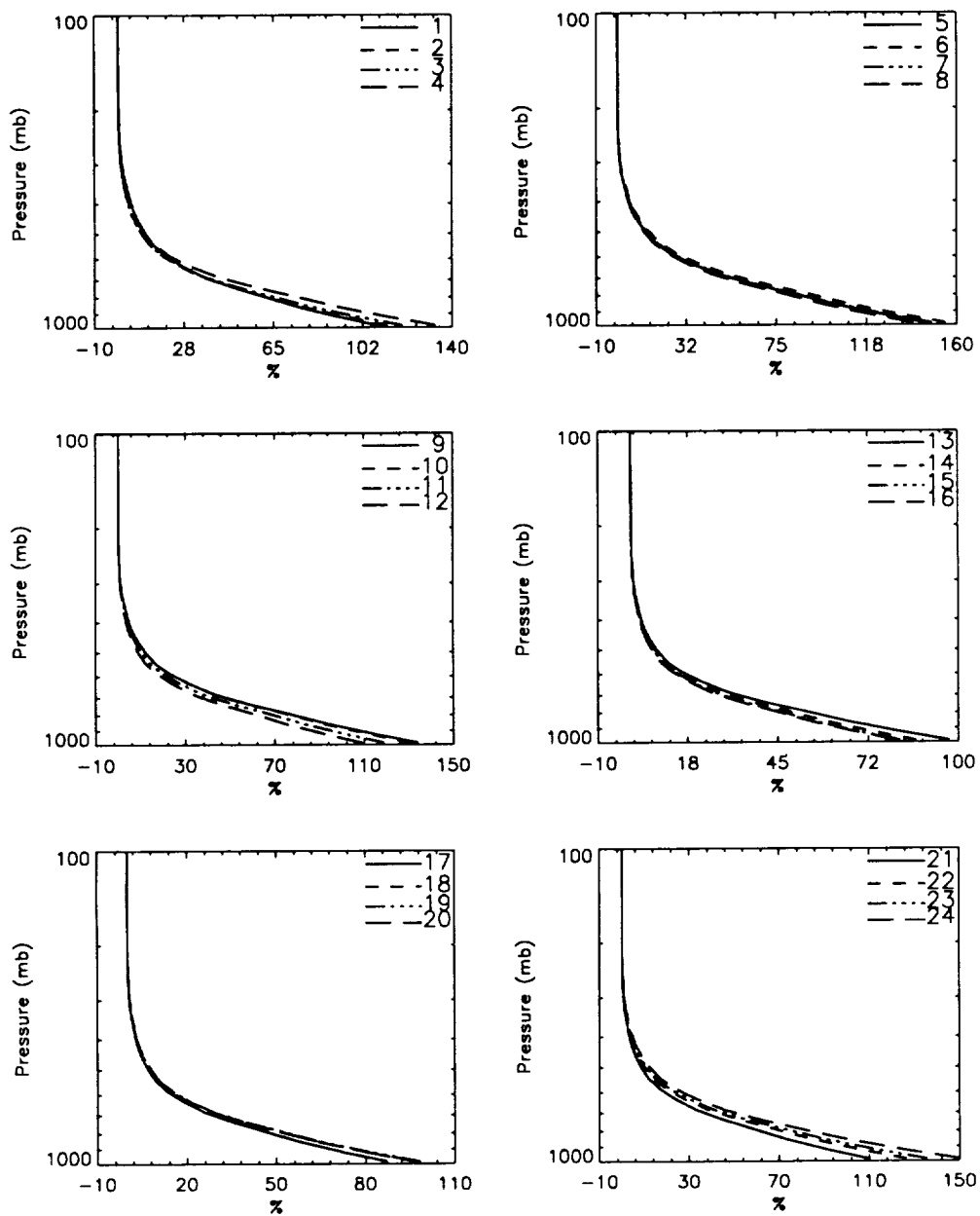


Figure 4.17: Effect of including interfering species on ECFs from selected channels computed for an ozone only atmosphere.

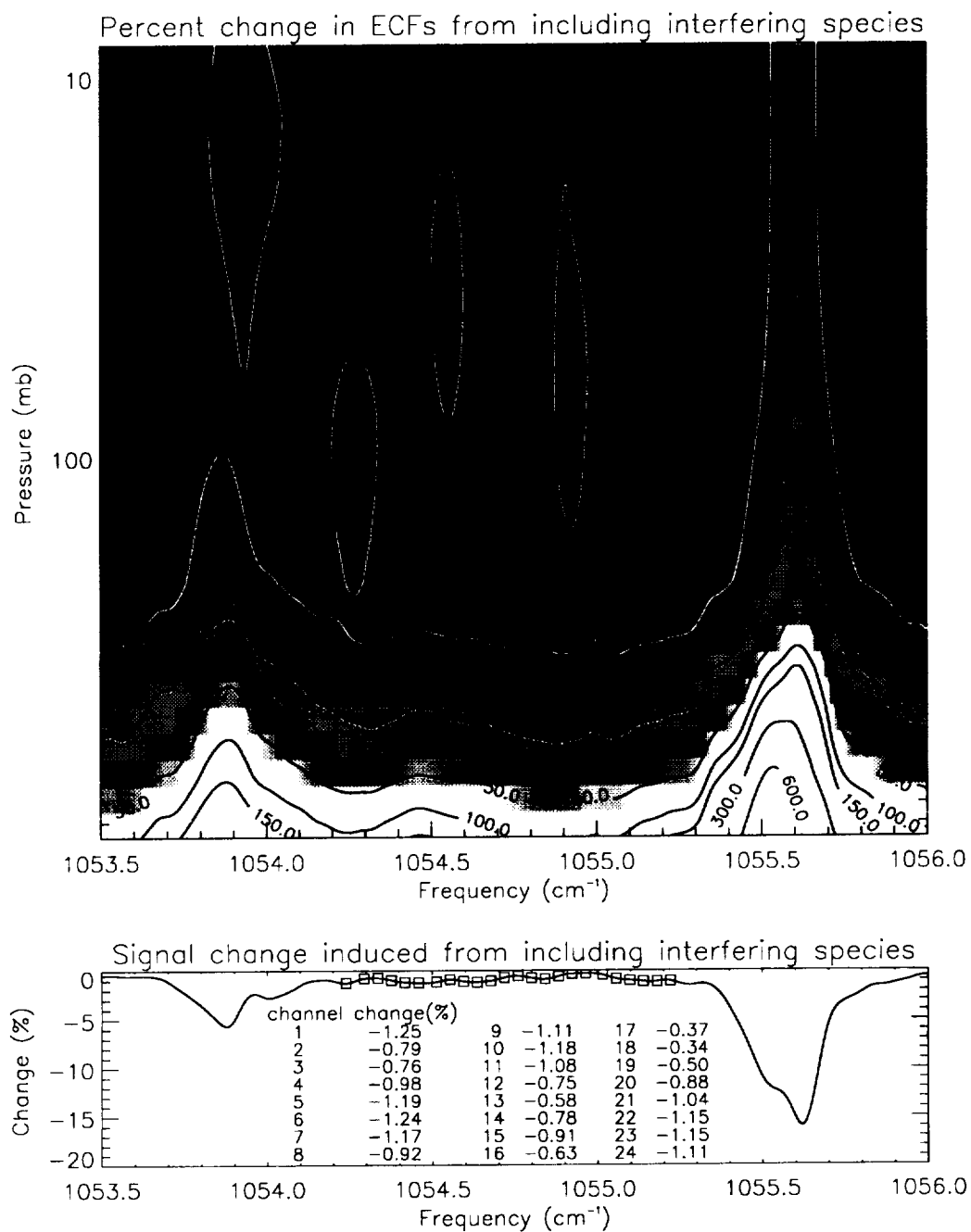


Figure 4.18: Effect on O_3 ECFs within the measurement band from including interfering species. Note that the channel locations attempt to minimize impact from the CO_2 (≈ 1053.92 and 1055.63 cm^{-1}) and H_2O ($\approx 1055.52 \text{ cm}^{-1}$) spectral lines and their associated wing effects.

tor elements within the linear detector array (see section 4.6). Since these detector elements have a finite size, detection design considerations impose a fundamental limit to the total number of channels possible within this region. Approximate channel locations of interest were defined such that the measurements would provide knowledge from regions of peak sensitivity to O_3 variations, maximum vertical discrimination capability by having a range of O_3 ECFs, and the ability to reproduce the signal spectrum as a continuous function within the band region. The design specifications of the detector array were then varied until a “best fit” of the desired channel locations was achieved. Figure 4.15 shows the spectral positions for the selected set of 24 channels within the plots depicting the signal sensitivities to atmospheric O_3 perturbations. The corresponding vertical distributions of O_3 information content are illustrated in Figure 4.16 with plots of O_3 ECFs for all channels; note that the O_3 ECFs have been normalized to their peak values to enable inter-channel comparisons. These contribution functions demonstrate the ability to de-couple tropospheric O_3 signal contributions from those of the stratosphere using the chosen channels. Based upon the broadness of these ECFs, ozone profiling capability does not appear too likely. However, some vertical discrimination ability should be attainable for the stratosphere along with integrated amounts for tropospheric O_3 . As will be discussed and demonstrated in Section 5.5, the vertical character of a retrieved O_3 profile for this measurement can be similar to that of the true O_3 profile under certain situations. Table 4.4 contains peak values for the O_3 ECFs as well as other channel specific information. As mentioned earlier, the H_2O vapor continuum provides the most significant impact from the interfering species within the instrument band region. Accordingly, interfering species primarily effect the O_3 ECFs in the tropospheric region; this is demonstrated in Figure 4.17 which emphasizes the lower

atmosphere. An attempt was made to minimize the potential impact of interfering species while positioning the instrument band. This can be appreciated upon viewing the effect on O_3 ECFs from including interfering species for a spectral interval slightly larger than the chosen channel region, as is depicted in Figure 4.18. Clearly, it is desirable to avoid overlap with the CO_2 (≈ 1053.92 and 1055.63 cm^{-1}) and H_2O ($\approx 1055.52 \text{ cm}^{-1}$) spectral lines and their associated wings. It should also be noted that the impact of interfering species would be equivalent to these calculations if and only if their existence was ignored; whereas, in reality the impact would be much less since significant knowledge pertaining to the interferants should exist. The potential effect from uncertainties in knowledge of the interferants will be addressed in the next chapter.

4.6 Further Design Aspects and Signal-to-Noise Ratio Analysis

The light gathering power of an optical system is characterized by its etendue or throughput, which can be given (Wyatt, 1991) by

$$U = A\Omega, \quad (4.23)$$

where A is the receiving area through some surface and Ω is a projected solid angle; and the projected solid angle can be defined as the area of another surface which is intercepted (or viewed) by the receiving surface divided by the square of the distance separating the surfaces. Figure 4.19 presents a schematic of the instrument design configuration proposed for this research. Accordingly, the telescope throughput of this instrument can be represented by

$$U_T = \frac{\pi}{4} d_T^2 \frac{\pi}{4} \theta_{IFOV}^2, \quad (4.24)$$

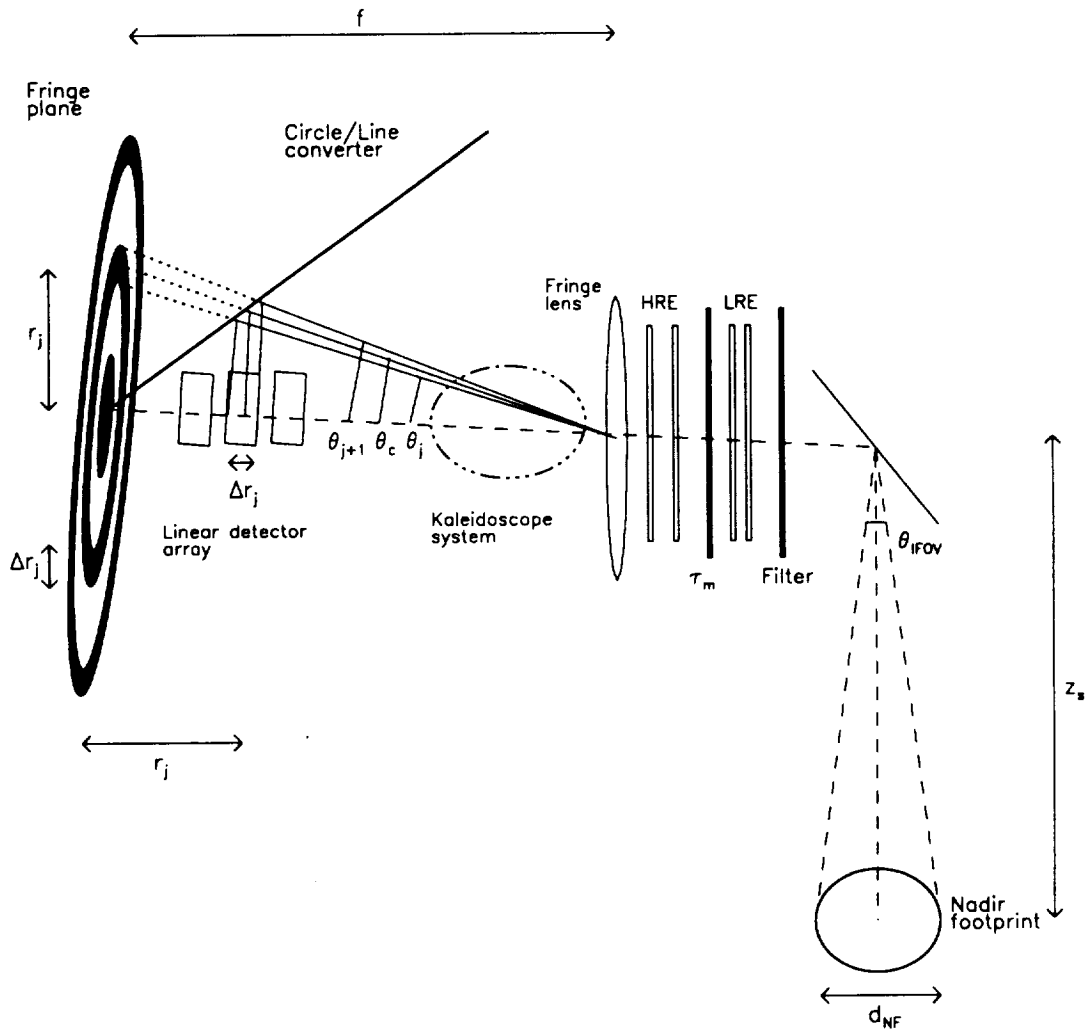


Figure 4.19: Schematic of proposed double-etalon instrument design configuration illustrating relative positioning of key optical elements.

with the instrument's instantaneous field-of-view angle (θ_{IFOV}) obtainable from

$$\tan\left(\frac{\theta_{IFOV}}{2}\right) = \frac{d_{NF}}{2z_s}; \quad (4.25)$$

d_T is the telescope entrance diameter, d_{NF} is the nadir footprint diameter, and z_s corresponds to the satellite altitude. Similarly, an expression for the FPI throughput can be written as

$$U_{FPI} = \frac{\pi}{4} D_E^2 \Omega_{FPI}, \quad (4.26)$$

with the Fabry-Perot interferometer solid angle Ω_{FPI} represented by

$$\Omega_{FPI} = \pi \tan^2 \theta_{max}; \quad (4.27)$$

D_E is the diameter of the etalons, and θ_{max} is the angle from the optical axis to the most outside fringe detected, which is determined by positioning of the linear array detector element farthest from the cone vertex. The angle θ_{max} can be given by

$$\tan \theta_{max} = \frac{r_{max}}{f}, \quad (4.28)$$

for fringe detection at a maximum distance of r_{max} from the cone vertex by an optical system having a fringe forming lens of focal length f . The projected solid angle can also be represented by $\Omega = \pi \sin^2 \alpha$, where α is the beam convergence angle with respect to the optical axis at the throughput evaluation surface; this is a more useful expression when evaluating channel throughput. By assuming invariance of throughput for the optical system and evaluating the throughput at various points along the beam path, several key optical design parameters may be calculated. For example, by specifying the nadir footprint size, satellite orbit altitude, and telescope diameter, the telescope throughput may be computed using equation 4.24. This result combined with the invariance of throughput requirement can then be used with equation 4.26 to calculate the necessary diameter of the etalon working area.

However, evaluation of Ω_{FPI} necessitates knowledge of r_{max} which requires specification of the spectral extent of the measurement. Expressions relating these quantities can be developed by considering in further detail the geometry of the CLIO system configuration. Upon viewing the instrument's optical and detection configuration in Figure 4.19, the following relationships can be realized:

$$\tan \theta_c = \frac{r_j}{f}, \quad (4.29)$$

$$\tan \theta_j = \frac{r_j - \frac{\Delta r_j}{2}}{f}, \quad (4.30)$$

$$\tan \theta_{j+1} = \frac{r_j + \frac{\Delta r_j}{2}}{f}; \quad (4.31)$$

θ_c , θ_j , and θ_{j+1} correspond to the angles between the optical axis and the detector center, starting, and ending positions, respectively; r_j is the radial distance from the cone vertex to the detector center along the optical axis (or to the detector's central spectral position in the fringe plane), Δr_j is the detector spatial length, and f is the focal length of the fringe forming lens as previously mentioned. Using equation A.3 we can write an expression relating the observation wavenumber and angle from the optical axis as

$$\nu = \frac{n}{2\mu t \cos \theta} = \frac{\nu_0'}{\cos \theta}, \quad (4.32)$$

where ν_0' is the wavenumber for $\theta = 0^\circ$. This equation enables a more useful relationship for θ_{max} to be formulated as

$$\cos \theta_{max} = \frac{\nu_0'}{\nu_0' + \Delta \nu_{obs}}, \quad (4.33)$$

where $\Delta \nu_{obs}$ is the desired spectral extent of the measurement. Equation 4.32 also permits development of an expression for the wavenumber extent (or spectral resolution) of the linear detector array elements (or channels) to be given by

$$\Delta \nu_L = \nu_0' \left[\frac{1}{\cos \theta_{j+1}} - \frac{1}{\cos \theta_j} \right]. \quad (4.34)$$

As can be seen from the above equations, the lengths (in the optical axis direction) of the detector array elements determine the spectral resolutions of the instrument channels. Whereas, the minimum necessary detector widths are determined by the detector plane spatial energy distribution perpendicular to the optical axis resulting from the CLIO system. The 90° kaleidoscope mirrors fold the fringe image (formed by the fringe forming lens) on top of itself, thus reducing the image area by a factor of 4. As discussed in Section 3.3, the view from the fringe plane is composed of 3 virtual images along with the real image (see Figure 3.7); this increases the etalon effective diameter to that given by equation 3.5. As with throughput, the relative aperture or f-number ($F_\#$) is a measure of the flux-gathering ability of an optical system. The f-number is given by the ratio of the effective focal length to the entrance aperture diameter; for this instrument we can write

$$F_\# = \frac{f}{D_E}, \quad (4.35)$$

for the FPI f-number. However, the kaleidoscope mirrors cause the “system f-number” to be reduced since $D_{eff} > D_E$, as given by

$$F_{\#,y} = \frac{f}{D_{eff}}. \quad (4.36)$$

A smaller f-number usually implies a better system, however, the associated cost and complexity level usually increase as well. An $F_{\#,y} = .5$ can be considered as a theoretical minimum, while $F_{\#,y} < 1$ is generally difficult to achieve in practice. In addition, the expressions for throughput given above become less accurate for lower f-number systems; however, they are sufficient for most feasibility studies and will be used in the present study with $F_{\#,y} \approx 1.2$ set as a design goal.

Following the kaleidoscope, the circle to line converter converts a 90° arc of the concentrated fringe image into a line directly below in the detector plane, while

leaving the radial (or spectral) direction unchanged. The length of this “line” which varies with distance from the cone vertex is the minimum width for the detector array element at a particular location and can be determined by assuming invariance of throughput. As indirectly mentioned earlier, the invariance of throughput can be stated as $A\pi \sin^2 \alpha = \text{constant}$. We can apply this condition at the “imaginary” fringe plane (which would exist in absence of the circle to line converter) and at the plane containing the detector array elements. For a particular channel in the fringe plane, the area can be given by $A_f = \frac{\pi f^2 [\tan^2 \theta_{j+1} - \tan^2 \theta_j]}{4}$, where the factor of 4 accounts for the kaleidoscope impact; in addition, $\sin^2 \alpha$ can be approximated by $\frac{1}{4F_{\#ys}^2}$. Alternatively, the area of the rectangular detector array elements in the detector plane can be represented by $A_d = \Delta r_j \mathcal{W}_j$, where Δr_j is the length and \mathcal{W}_j is the width; and $\alpha = \frac{\pi}{4}$ represents the half cone angle. Equating these throughputs and solving for the minimum necessary width, one obtains

$$\mathcal{W}_j = \frac{\pi f^2 [\tan^2 \theta_{j+1} - \tan^2 \theta_j]}{8F_{\#ys}^2 \Delta r_j}. \quad (4.37)$$

The diffraction limitation must be considered in the design process for any optical system to ensure sufficient image quality. The quality of an optical system is usually expressed in terms of image size for a point source. The image of a point source is not a point but is usually referred to as an optical blur. The blur circle for a perfect optical system is called an “Airy disk,” and is composed of a central blur of light surrounded by alternating dark and light rings of rapidly decreasing intensity. The size of the blur circle is usually defined by the diameter of the first dark ring as

$$D_{blur} = 2.44\lambda \frac{f}{D_{eff}} = 2.44\lambda F_{\#ys}, \quad (4.38)$$

where λ is the wavelength of radiation; f , D_{eff} and $F_{\#ys}$ are the lens focal length, effective system diameter and f-number, respectively, as previously defined. The

Characteristics of Assumed Orbit	
Orbit:	Sun-synchronous; altitude $z_s = 955$ km
Scan:	$\pm 51^\circ$ from nadir
Scan rate (s):	$\frac{1 \text{ scan}}{8 \text{ sec}} = .125 \frac{\text{scan}}{\text{sec}}$
Instantaneous field-of-view (IFOV):	3° ; 50 km at nadir
#pixels / scan (n):	$1 + \frac{(2 \times 51^\circ)}{3^\circ} = 35$
Step & dwell time:	$t_{sd} = \frac{1}{ns} = 228.57$ ms
Dwell (integration) time:	$t_d \approx 220$ ms

Table 4.5: Orbital characteristics and dwell time determination assuming synchronization with the Nimbus-7/TOMS orbit.

angular size of the blur can be given by the ratio of the blur diameter to the effective focal length as

$$\theta_{blur} = \frac{2.44\lambda}{D_{eff}}. \quad (4.39)$$

In order to minimize the loss of energy due to diffraction, the angular width of the narrowest fringe

$$\Delta\theta_{f_{min}} = [\theta_{j+1} - \theta_j]_{min} \approx \frac{[\Delta r_j]_{min}}{f} \quad (4.40)$$

must be larger than the blur circle. Therefore, the condition that $\frac{\Delta\theta_{f_{min}}}{\theta_{blur}} > 1$ must be verified in the design process.

Adequate signal-to-noise ratio (SNR) is essential for the successful design of any electro-optical system intended for information processing. The SNR can be evaluated from the radiometric performance equation, which for a uniform extended-area source can be expressed (Wyatt, 1991) as

$$SNR = \frac{D^* \tau_i A_s \Omega_s}{\beta \sqrt{A_d \Delta f}} \int_{\nu} I(\nu) \mathcal{R}_i(\nu) d\nu; \quad (4.41)$$

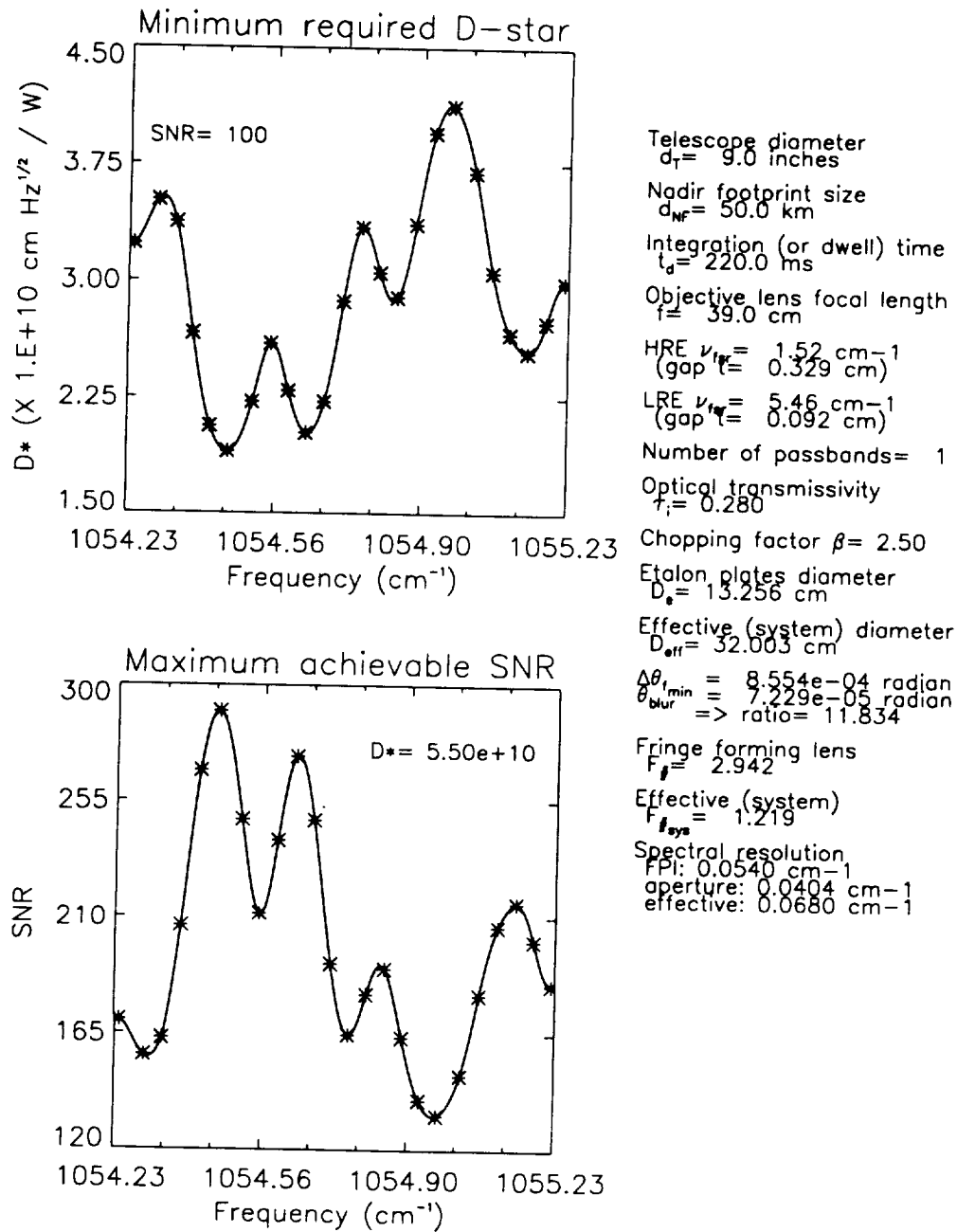


Figure 4.20: Results from a signal-to-noise ratio (SNR) analysis using the U.S. Standard (1976) atmosphere. Spectral locations for channels are indicated by *.

Channel	$\nu_{0_c}(cm^{-1})$	$\theta(radian)$	$r(cm)$	$\Delta r(cm)$	$W(cm)$	SNR
1	1054.240	1.07e-02	0.42	1.42e-01	2.17e-01	169.97
2	1054.295	1.48e-02	0.58	1.01e-01	3.03e-01	156.44
3	1054.336	1.72e-02	0.67	8.70e-02	3.54e-01	162.99
4	1054.377	1.93e-02	0.75	7.73e-02	3.98e-01	206.50
5	1054.418	2.12e-02	0.83	7.03e-02	4.38e-01	266.90
6	1054.459	2.30e-02	0.90	6.49e-02	4.74e-01	289.88
7	1054.513	2.51e-02	0.98	5.94e-02	5.18e-01	247.79
8	1054.554	2.66e-02	1.04	5.61e-02	5.49e-01	211.52
9	1054.595	2.81e-02	1.09	5.32e-02	5.79e-01	239.65
10	1054.636	2.94e-02	1.15	5.08e-02	6.07e-01	272.47
11	1054.677	3.07e-02	1.20	4.86e-02	6.33e-01	247.51
12	1054.718	3.19e-02	1.25	4.68e-02	6.59e-01	191.67
13	1054.759	3.31e-02	1.29	4.51e-02	6.84e-01	164.57
14	1054.800	3.43e-02	1.34	4.36e-02	7.07e-01	180.02
15	1054.841	3.54e-02	1.38	4.22e-02	7.30e-01	189.86
16	1054.882	3.65e-02	1.42	4.09e-02	7.53e-01	163.44
17	1054.923	3.75e-02	1.46	3.98e-02	7.74e-01	139.08
18	1054.964	3.86e-02	1.50	3.87e-02	7.96e-01	133.20
19	1055.016	3.98e-02	1.55	3.75e-02	8.22e-01	148.72
20	1055.057	4.08e-02	1.59	3.66e-02	8.41e-01	179.91
21	1055.098	4.17e-02	1.63	3.58e-02	8.61e-01	206.16
22	1055.139	4.26e-02	1.66	3.50e-02	8.80e-01	215.83
23	1055.180	4.35e-02	1.70	3.43e-02	8.99e-01	200.74
24	1055.221	4.44e-02	1.73	3.36e-02	9.17e-01	183.83

Table 4.6: Summary of detector element positioning and achievable SNR (assuming $D^* = 5.5 \times 10^{10} cm Hz^{\frac{1}{2}}/W$) for the chosen channels.

D^* (D-star) is the detectivity, in $cmHz^{\frac{1}{2}}/W$; τ_i is the instrument peak optical transmittance; A_s and Ω_s are the area (cm^2) and solid angle (sr), respectively, of the source, so that $A_s\Omega_s$ represents the source throughput; β is the chopping factor; A_d ($= \Delta r_j \mathcal{W}_j$) is the detector area, in cm^2 ; Δf is the electrical information-noise bandwidth, in s^{-1} ; $I(\nu)$ is the upwelling radiance ($Wcm^{-2}sr^{-1}/cm^{-1}$) and $\mathcal{R}_i(\nu)$ is the instrument relative spectral response, so that $\int_{\nu} I(\nu)\mathcal{R}_i(\nu)d\nu$ represents the effective flux, in $Wcm^{-2}sr^{-1}$. Applying equation 4.41 to this instrument feasibility study, we can replace $\int_{\nu} I(\nu)\mathcal{R}_i(\nu)d\nu$ with $S(\nu)$ (equation 4.19), the FPI signal. By invariance of throughput, the source throughput $A_s\Omega_s$ can be replaced by the channel throughput $U(\nu) = A_d(\nu)\pi\sin^2(\frac{\pi}{4})$. In addition, the information-noise bandwidth may be expressed (Wyatt, 1991) as $\Delta f = \frac{5}{2\pi t_d}$, where t_d is the scene integration or dwell time. We can now reformulate the radiometric performance equation (equation 4.41) for the feasibility study in this research as

$$SNR(\nu) = \frac{D^* \tau_i A_d(\nu) \pi \sin^2(\frac{\pi}{4})}{\beta \sqrt{A_d(\nu) \frac{5}{2\pi t_d}}} \int_{\nu_{low}}^{\nu_{high}} T_D(\nu - \nu') T_F(\nu') I(\nu') d\nu'. \quad (4.42)$$

At the expense of slightly additional mechanical complexity, the detectivity is desired to be as high as reasonably possible to enable achieving maximum SNR; accordingly, this design will require detector cooling for improved detectivity, in addition to cold stop limitation of unnecessary portions of the detector element FOV's to block unwanted background radiation from being detected. Mercury Cadmium Telluride (HgCdTe) photoconductive (PC) detectors perform quite well in this spectral region; cryogenic operation at 77 K utilizing liquid nitrogen (LN_2) cooling and 60° FOV cold stop shielding will be assumed in this work. A detectivity value of $D^* = 5.5 \times 10^{10} cmHz^{\frac{1}{2}}/W$ will be implemented into the feasibility study calculations. Many detector manufacturers can meet or exceed this D-star value (i.e., the J15D series by EG&G Judson) for the narrow bandwidth, small detector element, limited

FOV, linear detector array configuration being considered for this $10\text{ }\mu\text{m}$ region. A chopping factor of $\beta = 2.5$ will be used. The dwell time is defined as the time that the system dwells on a single resolution element and can be given approximately by $t_d \approx \frac{1}{ns}$, where n is the number of resolution elements per scan and s is the scan rate. Assuming a 50km nadir footprint and orbital parameters similar to those of the Total Ozone Mapping Spectrometer (TOMS) instrument aboard Nimbus-7 (Greg et al., 1991) to achieve continuous coverage at the equator, yields a dwell time of $t_d \approx 220\text{ ms}$. The orbital characteristics and dwell time determination for this study are summarized in Table 4.5. Figure 4.20 represents the results of this SNR analysis. The top plot depicts the minimum D^* necessary to achieve a SNR=100. Whereas, the maximum achievable SNR for $D^* = 5.5 \times 10^{10} \text{ cmHz}^{\frac{1}{2}}/W$ is illustrated in the bottom plot; these values will be assumed for the error analysis of the next chapter. Table 4.6 summarizes the design logistics for the chosen channels. Note that all suggested instrument design specifications are contained within Tables 4.3 and 4.6 as well as Figure 4.20.

CHAPTER V

RETRIEVAL SIMULATION AND ERROR ANALYSIS

5.1 Introduction

Retrieval of atmospheric ozone will be performed through implementation of a nonlinear least squares iterative technique to achieve convergence of the “measured” and “calculated” forward model signals. In particular, the maximum likelihood method will be implemented in an approach similar to that discussed by Rodgers (1976, 1984); this is a statistically based technique which yields the most probable solution. For the error analysis, a procedure similar to Rodgers (1989, 1990) will be implemented.

5.2 Ozone Retrieval Technique

Retrieval theory enables the derivation of atmospheric parameters from measured values of radiation emerging from the top of the atmosphere, the intensity and spectral distribution of which depend on the state of the atmosphere according to known physics. The task is to deduce the best estimate of the state of the atmosphere or the particular parameters of interest from the radiation measurements. As discussed by Rodgers (1976), there are two distinct aspects to this problem, namely the “inverse”

problem and the “estimation” problem. The inverse problem involves inverting a known equation that expresses the measured radiation as a function of the atmospheric state, into a formulation for the atmospheric state in terms of the radiation; however, this result usually has no mathematically unique solution and is said to be “ill-conditioned.” Nonuniqueness in the solution arises from several factors. There are vertical regions of the atmosphere which provide minimal or no contribution to the measured upwelling radiances, and therefore their corresponding constituent concentrations could be limitless; a priori constraints can be used to impose boundaries on their sizes. In addition, measurement noise in the observations implies a corresponding nonuniqueness in the solution; the magnitude of which can be estimated by mapping measurement error bars onto error bars of the solution. Nonuniqueness is also introduced into the solution from a finite vertical resolution, which results from a finite spectral resolution of the instrument as well as a consequence of the nadir-viewing geometry; note that even with infinite spectral resolution, finite vertical resolution would still exist when employing a nadir-viewing satellite configuration. The process of determining the best solution from all the possible ones constitutes an estimation problem.

A constrained least squares solution, such that the sum of the squares of the differences between the measurements and the calculated signal values corresponding to the solution is minimized, is commonly implemented to solve an ill-conditioned problem. A priori constraints can refer to any type of constraints on a solution and they can be regarded as “virtual measurements” (Rodgers, 1976), since they contain information pertaining to the unknown profile similarly to the direct measurements. If nothing is known about the unknown function, one can obtain an inversion using some sort of objective constraint, such as minimization of variance (Twomey, 1977).

If statistics are known, then one can constrain accordingly. For example, from a statistical constraint approach, a climatological mean profile and its covariance can be considered as a measurement and its corresponding uncertainty; and a constrained solution can be found which minimizes the mean square departure from this mean. This method tends to push the indirectly sensed solution towards conformity with a body of past data possibly obtained by more direct methods. This formula is very useful when there is justification for choosing the statistical constraint.

The atmospheric remote sounding retrieval problem is usually nonlinear; consider, for example, the extraction of an atmospheric constituent vertical profile from equation 4.14. In general, we want to know the value of an unknown vector \mathbf{x} describing the state of a particular atmospheric parameter; this is usually a vertical profile of some quantity given at a finite number n of levels (i.e., the O_3 mixing ratio profile). We measure a signal vector \mathbf{y} containing m elements and possessing an $m \times m$ error covariance matrix \mathbf{S}_e , which is related to \mathbf{x} by

$$\mathbf{y} = F(\mathbf{x}), \quad (5.1)$$

where F is a nonlinear forward model. Often $m \ll n$, making the inverse problem formally ill-posed; as is the case for this study with $m = 24$ (the number of instrument channels) and $n = 46$ (the number of levels used to represent the vertical O_3 profile). Assuming we have a priori information \mathbf{x}_0 about \mathbf{x} with an $n \times n$ error covariance matrix \mathbf{S}_x , the maximum likelihood condition (Rodgers, 1976) requires that \mathbf{x} minimizes

$$[F(\mathbf{x}) - \mathbf{y}]^T \mathbf{S}_e^{-1} [F(\mathbf{x}) - \mathbf{y}] + [\mathbf{x} - \mathbf{x}_0]^T \mathbf{S}_x^{-1} [\mathbf{x} - \mathbf{x}_0]. \quad (5.2)$$

Although usually difficult to implement, the solution can be found by finding \mathbf{x} that minimizes equation 5.2. However, easier approaches are usually available if

the problem is not too nonlinear, or if good a priori information is available in the very nonlinear case. For example, nonlinear problems are frequently solved with some form of Newtonian iteration, which requires the forward model to be expanded as a Taylor series about a guessed value \mathbf{x}_i of the solution. This enables a direct measurement of $\mathbf{y} = F(\mathbf{x})$ to be expressed as

$$\mathbf{y} = F(\mathbf{x}_i) + (\mathbf{x} - \mathbf{x}_i) \frac{\partial F}{\partial \mathbf{x}} \big|_{\mathbf{x}_i}, \quad (5.3)$$

where higher-order terms have been neglected by using a truncated Taylor series. This may be reformulated as

$$\mathbf{y} = \mathbf{y}_i + \mathbf{K}_{x_i}(\mathbf{x} - \mathbf{x}_i), \quad (5.4)$$

through defining the calculated signal for \mathbf{x}_i as $\mathbf{y}_i = F(\mathbf{x}_i)$ and the forward model derivative by

$$\mathbf{K}_{x_i} = \frac{\partial F}{\partial \mathbf{x}} \big|_{\mathbf{x}_i}; \quad (5.5)$$

\mathbf{K}_x is an $m \times n$ matrix which is sometimes called the weighting function. This results in an under-constrained set of equations for \mathbf{x} . Considering the a priori constraint \mathbf{x}_0 with covariance \mathbf{S}_x as a virtual measurement yields the following set of equations:

$$\mathbf{K}_{x_i} \mathbf{x} = \mathbf{y} - \mathbf{y}_i + \mathbf{K}_{x_i} \mathbf{x}_i \quad \text{with} \quad \text{cov}(\mathbf{K}_{x_i} \mathbf{x}) = \mathbf{S}_\epsilon, \quad (5.6)$$

and

$$\mathbf{x} = \mathbf{x}_0 \quad \text{with} \quad \text{cov}(\mathbf{x}) = \mathbf{S}_x. \quad (5.7)$$

Two independent measurements of a scalar quantity are usually combined by taking a weighted average, using the reciprocal of the variances as weights. Using the same principle applied to vector quantities, the direct measurement and virtual measurement of equations 5.6 and 5.7, respectively, can be combined to get the best

estimate of \mathbf{x} (the O_3 mixing ratio profile). As shown by Rodgers (1976), this solution can be expressed by an iteration equation as

$$\mathbf{x}_{i+1} = \mathbf{x}_0 + \mathbf{S}_x \mathbf{K}_{x_i}^T (\mathbf{K}_{x_i} \mathbf{S}_x \mathbf{K}_{x_i}^T + \mathbf{S}_\epsilon)^{-1} [\mathbf{y} - \mathbf{y}_i - \mathbf{K}_{x_i}(\mathbf{x}_0 - \mathbf{x}_i)], \quad (5.8)$$

with a covariance of the solution given by the $n \times n$ matrix

$$\mathbf{S} = (\mathbf{S}_x^{-1} + \mathbf{K}_x^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_x)^{-1}. \quad (5.9)$$

Equation 5.8 can be implemented in a sequential fashion when applied to actual satellite data, whereby the new estimate of \mathbf{x} made at each measurement stage is used as the a priori vector for the next stage. The concept of sequential estimation allows one to use the continuity that exists along a subsatellite track to improve the accuracy of retrieval. However, in practice it can be difficult to estimate \mathbf{S}_x . If knowledge of how the profile should change horizontally is available (i.e., through analysis of global climatology), then this can be combined with the retrieval previous in time and therefore close in space to formulate a better a priori estimate using an extrapolative process. A best estimate of \mathbf{S}_x can be obtained by a statistical analysis of the a priori data set, for example, \mathbf{S}_x can be defined to be the statistical variance of the climatological average profile at the horizontal point of interest. As is obvious from equations 5.8 and 5.9, changing \mathbf{S}_x due to uncertainty in its appropriate value directly affects the estimated solution \mathbf{x} and its corresponding covariance \mathbf{S} .

5.3 Method of Error Analysis

A procedure similar to that discussed by Rodgers (1989, 1990) will be followed for the analysis of error in this research. The total error can be separated into components due to instrument measurement noise, portions of the vertical profile which cannot be seen by the observing system, and uncertainty in model parameters.

The “observing system” refers to a combination of both the measuring instrument and the retrieval technique; hence an actual atmospheric constituent profile is the input to the observing system, while a retrieved profile is its output.

The measured signal vector in equation 5.1 can be re-expressed as

$$\mathbf{y} = F(\mathbf{x}, \mathbf{b}) + \epsilon_y, \quad (5.10)$$

where \mathbf{b} is a vector of forward model parameters (such as instrument calibration characteristics, spectral line data, surface characteristics, etc.) which are not perfectly known, and the error vector ϵ_y represents measurement uncertainty. The uncertainty in the forward model parameters can be a source of random or systematic differences between measured and calculated values of \mathbf{y} . It is assumed that the forward model accurately represents the physics of the measurement process, so that forward model errors are due solely to uncertainty in forward model parameters; note that these model parameters can include quantities that are closely coupled to the state vector \mathbf{x} itself (i.e., spectral line parameters of the target species) but are not being deduced in the retrieval process.

The measurement error ϵ_y directly contributes to error $\mathbf{D}_y \epsilon_y$ in the retrieval process, where \mathbf{D}_y is an $n \times m$ matrix representing the contribution functions and may be written (Rodgers, 1976) as

$$\mathbf{D}_y = (\mathbf{S}_x^{-1} + \mathbf{K}_x^T \mathbf{S}_\epsilon^{-1} \mathbf{K}_x)^{-1} \mathbf{K}_x^T \mathbf{S}_\epsilon^{-1}. \quad (5.11)$$

This corresponds to an error covariance matrix of

$$\mathbf{S}_m = \mathbf{D}_y \mathbf{S}_\epsilon \mathbf{D}_y^T. \quad (5.12)$$

If the measurement channels (or elements of the signal vector \mathbf{y}) are independent, then \mathbf{S}_ϵ is a diagonal matrix. As defined by Rodgers (1990), the “null-space error”

results from regions of profile space which cannot be measured by the observing system. Assuming the mean profile to be used for \mathbf{x}_0 as an a priori constraint, the null-space error for any single retrieved profile can be given by $(\mathbf{A}_x - \mathbf{I})(\mathbf{x} - \bar{\mathbf{x}})$; \mathbf{A}_x is the averaging kernel matrix defined by $\mathbf{A}_x = \mathbf{D}_y \mathbf{K}_x$, \mathbf{I} is the identity matrix, and $\bar{\mathbf{x}}$ represents the mean state so that $\mathbf{x} - \bar{\mathbf{x}}$ corresponds to the true profile's deviation from the mean state. Knowledge of the statistics of $\mathbf{x} - \bar{\mathbf{x}}$ is necessary to reasonably estimate the null-space error; and, in particular, a covariance matrix \mathbf{S}_x of its variability is required. The null-space error can then be expressed by

$$\mathbf{S}_n = (\mathbf{A}_x - \mathbf{I})\mathbf{S}_x(\mathbf{A}_x - \mathbf{I})^T. \quad (5.13)$$

As previously mentioned, errors $\epsilon_b = \mathbf{b} - \hat{\mathbf{b}}$ in the forward model parameters can be random or systematic in nature, depending on their source of origination. A vector $\hat{\mathbf{b}}$ can be defined as a best estimate of the model parameters, having an error covariance matrix \mathbf{S}_b , which is used in the retrieval process; whereas, the measuring system uses exact values for the model parameters \mathbf{b} . \mathbf{S}_b is an $nb \times nb$ matrix, where nb is the number of uncertain forward model parameters which are being considered as contributors to the overall forward model error; in addition, the vectors \mathbf{b} and $\hat{\mathbf{b}}$ both contain nb elements. This introduces a model parameter component into the retrieval error of $\mathbf{D}_y \mathbf{K}_b \epsilon_b$, with a covariance matrix given by

$$\mathbf{S}_s = \mathbf{A}_b \mathbf{S}_b \mathbf{A}_b^T; \quad (5.14)$$

the measurement sensitivity to forward model parameters is represented by the $m \times nb$ matrix \mathbf{K}_b and is defined by

$$\mathbf{K}_b = \frac{\partial F}{\partial \mathbf{b}}; \quad (5.15)$$

and the $n \times nb$ matrix $\mathbf{A}_b = \mathbf{D}_y \mathbf{K}_b$ corresponds to the state parameter sensitivity to model parameters \mathbf{b} . Note that \mathbf{S}_m , \mathbf{S}_n , and \mathbf{S}_s are all $n \times n$ covariance matrices. As

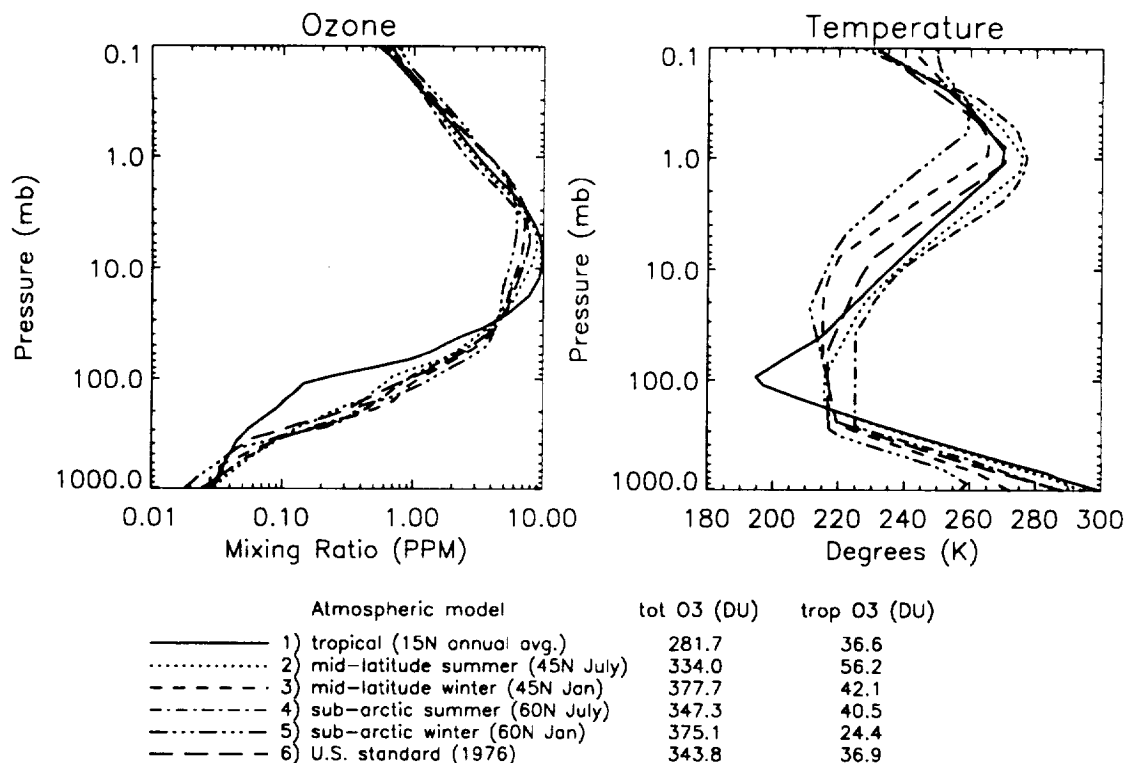


Figure 5.1: Atmospheric models from the 1976 Standard Atmosphere database.

discussed above, the overall error in the retrieved profiles results from a combination of three components: measurement error, null-space error, and forward model error. The total random error can then be formulated by summing the null-space and measurement errors, in addition to any random components of the forward model error.

5.4 Application to Atmospheric Cases

An atmospheric database containing temperature, pressure, density, and gas volume mixing ratios (Anderson et al., 1986) of the necessary chemical species as a function of altitude for six reference atmospheres will be utilized for the retrieval and error analysis simulations in this research. The vertical profiles were formulated by a combination of measurements and theoretical predictions, and dayside estimates

for diurnally varying species (i.e., O_3) were incorporated. The database includes only those twenty-eight minor and trace atmospheric gases which were part of the AFGL molecular line parameter compilation (Rothman et al., 1987) at the time of publication; note that the most recent HITRAN database (Rothman, 1993) contains line parameters for thirty-one molecular species, however, as mentioned in Section 4.3 these additional species are relatively insignificant for the radiative transfer within the $9.6 \mu\text{m}$ spectral region of the current study. The reference atmospheric model profiles consist of the following types of atmospheres: 1) tropical (15N annual average), 2) mid-latitude summer (45N July), 3) mid-latitude winter (45N January), 4) subarctic summer (60N July), 5) subarctic winter (60N January), and 6) U.S. standard (1976). Vertical profiles of temperatures and ozone volume mixing ratios for these atmospheric types are illustrated in Figure 5.1.

The diagonal elements of a covariance matrix are the error variances of the elements of the corresponding vector which it represents; whereas, the off-diagonal elements indicate inter-level correlations existing between the variances at different levels. If the a priori information vector \mathbf{x}_0 can be derived from an independent data set of high spatial resolution, then it should be possible to compute an accurate covariance matrix \mathbf{S}_x ; however, such a data set is not always available and other means must be explored to arrive upon a reasonable estimate for \mathbf{S}_x . Realistic values for the variances at each level (or diagonal elements of \mathbf{S}_x) are usually obtainable, whereas the inter-level covariances (or off-diagonal elements of \mathbf{S}_x) may be somewhat harder to find. If the inter-level covariances are not available, then the variances of the different levels can be assumed independent and the off-diagonal elements of \mathbf{S}_x set to zero; alternatively, these off-diagonal elements ($S_{ij}, i \neq j$) can be estimated. As discussed by Rodgers (1990), by assuming a vertical length scale l of correlations, a

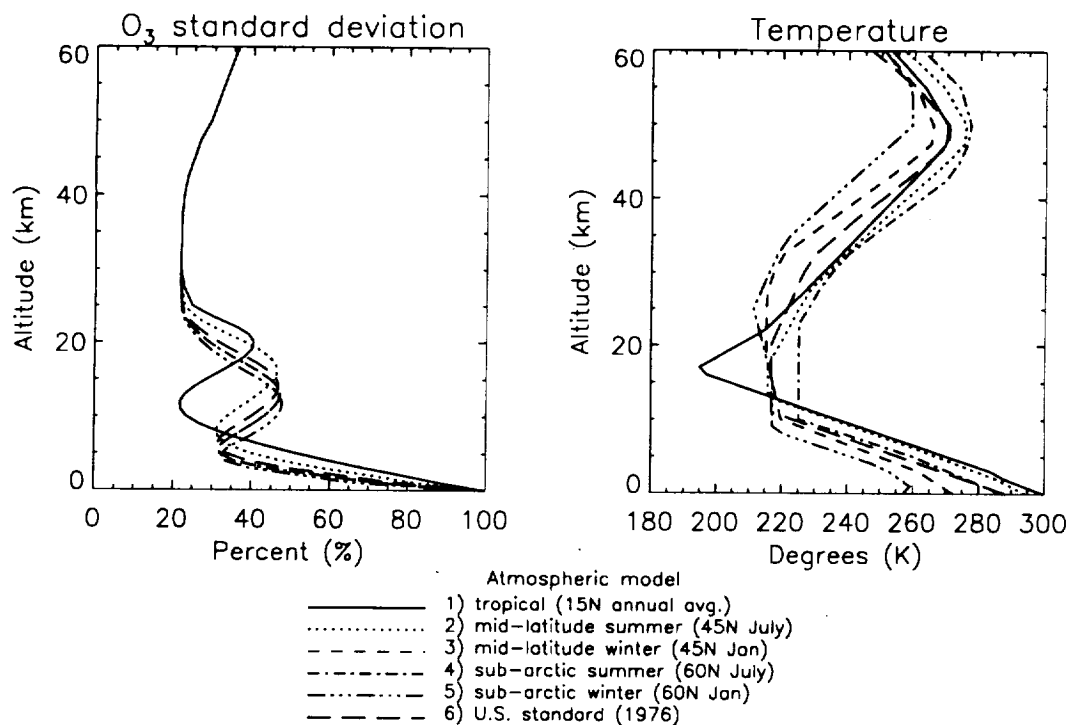


Figure 5.2: Ozone standard deviation profiles used in generating the a priori covariance matrices corresponding to atmospheric models within the 1976 Standard Atmosphere database. Note that atmospheric models 3) and 4) have similar tropopause heights (~ 12 km), making their corresponding standard deviation profiles overlap and indistinguishable (appearing as one very long-dashed profile).

formulation such as

$$S_{ij} = (S_{ii}S_{jj})^{\frac{1}{2}} \exp\left[\frac{-(z_i - z_j)^2}{l^2}\right] \quad (5.16)$$

can be used to generate the elements of an artificial covariance matrix; z_i and z_j correspond to altitudes (in the same units as l) of levels i and j , respectively. Equation 5.16 will be used to estimate the a priori covariance matrices in this study by assuming a vertical length scale of correlations $l = 5$ km for the O_3 covariances. The ozone variances at each level were arrived upon as follows. The uncertainty (in percent) of ozone concentrations in the Standard Atmosphere database (Anderson et al., 1986) is approximately constant throughout much of the stratosphere. At higher altitudes, the uncertainty increases primarily due to lack of measurements. And at altitudes below, local maxima exist in the lower stratosphere/upper troposphere and lower troposphere regions due to uncertainty in tropopause height and surface ozone variability, respectively. Based on these criteria, smooth curves were generated to approximate these ozone variances for use as diagonal elements in the a priori covariance matrices in the different atmospheric simulation cases. Figure 5.2 shows the resulting a priori ozone standard deviation profiles implemented within this analysis. Note the conservative approach used in the lower (than tropical atmosphere) tropopause atmospheric cases, where the standard deviations for lower atmosphere regions have been arbitrarily increased in an effort not to under estimate tropospheric errors for these cases possessing shallow tropospheres. It should be mentioned that while a more elaborate technique for S_x generation is certainly possible, the author believes that the method implemented enables the a priori uncertainty to be adequately represented for a feasibility study at this stage in an instrument design process.

The measurement channels will be assumed independent for this analysis. Ac-

Signal-to-noise ratio ranges		
Atmospheric model	SNR_{min}	SNR_{max}
1) tropical (15N annual average)	176.84	392.00
2) mid-latitude summer (45N July)	159.01	327.69
3) mid-latitude winter (45N Jan)	104.63	206.79
4) subarctic summer (60N July)	159.97	285.34
5) subarctic winter (60N Jan)	89.95	158.23
6) U.S. standard (1976)	133.20	289.88

Table 5.1: Summary of signal-to-noise ratio range estimates corresponding to the instrumentation of this research applied to different atmospheric types. Values of SNR for each spectral channel are necessary for formulation of the measurement covariance matrices.

cordingly, the measurement covariance matrix S_e will be a diagonal matrix with elements equal to the variance of the corresponding channels which they represent. Assuming the measurement error to be normally distributed and having a zero mean, the diagonal elements of the measurement covariance matrix can be estimated for each channel i by

$$S_{eii} = \left(\frac{1}{SNR_i}\right)^2 y_i^2, \quad (5.17)$$

in units of the measured signal squared (y_i^2); the signal-to-noise ratio for channel i is represented by SNR_i . Values of SNR_i , y_i , and thus S_{eii} were adjusted to correspond to the particular type of atmosphere being simulated. Table 5.1 summarizes the range in signal-to-noise ratios corresponding to the six atmospheres simulated in this research project; see Figure 4.20 for the spectral dependence of the SNR pertaining to the U.S. standard (1976) atmosphere.

The partial derivatives composing \mathbf{K}_x and \mathbf{K}_b (equations 5.5 and 5.15, respectively) were evaluated numerically by perturbing the forward model and using finite forward differences on the results. For example, all rows in column p of \mathbf{K}_x can be generated from

$$\mathbf{K}_{x_{(i=1-24, j=p)}} = \frac{F(\mathbf{x}_0 + \delta\mathbf{x}_0) - F(\mathbf{x}_0)}{\delta x_{0p}}, \quad (5.18)$$

where $\delta\mathbf{x}_0$ represents the perturbation to \mathbf{x}_0 and δx_{0p} corresponds to the perturbation magnitude at level p (the only non-zero element in $\delta\mathbf{x}_0$). The matrix \mathbf{K}_x is necessary for both the retrieval process and error analysis, whereas \mathbf{K}_b is needed only for the error analysis. In the error analysis, \mathbf{K}_x and \mathbf{K}_b are required to be evaluated only once for each atmospheric state about which the error budget is being considered. Whereas \mathbf{K}_x could be updated at each stage of the iteration process for each retrieval, if desired; although this step is rarely necessary, since convergence can usually be achieved in just a few iterations without derivative updates by using a reasonable first guess. As mentioned earlier, the vector \mathbf{b} (needed for \mathbf{K}_b evaluation) includes those forward model parameters for which the effect of their uncertainty on the retrieved ozone amounts is desired. It is beyond the scope of this present study to include all possible parameters, however, an attempt has been made to consider those most important. In particular, contributions to the overall error will be assessed (in Section 5.6) for the following forward model parameters: water vapor continuum model coefficients, surface temperature and emissivity, instrument radiometric channel calibrations, vertical profiles of atmospheric temperature and interfering species, and spectral line parameters.

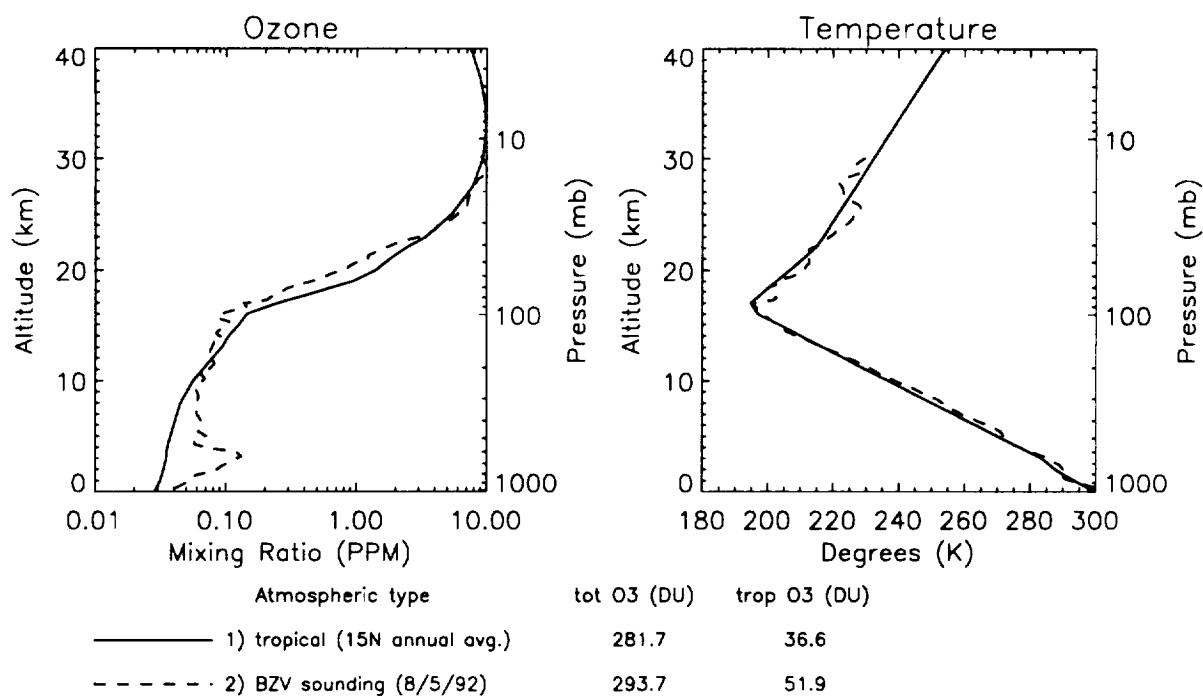


Figure 5.3: Ozone and temperature profiles from the 1976 Standard Atmosphere (tropical) along with a sounding from Brazzaville, Congo, on August 5, 1992.

5.5 Retrieval of O_3 for a Perturbed Tropical Atmosphere

The retrieval technique implemented into this study incorporates measurement uncertainty and null-space error for one retrieved profile, whereas the error analysis estimates how good the retrieved profiles will be in a statistical sense (encompassing many such retrievals) while including the effects of measurement, null-space, and forward model errors. Additionally, the mechanics of the retrieval process are the same for different atmospheric situations; therefore, the retrieval ability will be demonstrated for only one case, a perturbed tropical atmosphere. A tropical atmosphere has been chosen due to the significant recent interest in the south Atlantic ocean region where elevated tropospheric ozone has been observed (Fishman et al., 1990).

Sounding data from Brazzaville, Congo, on August 5, 1992, provided by Dr. Jack Fishman (NASA Langley Research Center) and Mr. Vincent Brackett (SAIC Corporation, Hampton, Virginia), were used for generation of the “true” profile. Although the ozone sounding profile was at high vertical resolution, the vertical levels did not coincide exactly with all of those needed for input into the forward model; the necessary interpolation was handled using cubic spline interpolation (see, for example, Press et al., 1990). In addition, the sounding profile stops below 11 mb (~ 30 km). Specification of the true O_3 profile for altitudes above 30 km was performed by using a modified average tropical atmosphere profile, and connecting the two profiles with cubic spline interpolation; the final modified profile was then formulated by adjusting the average tropical atmosphere profile with a multiplicative factor until the total O_3 column abundance matched the value provided with the sounding data (which was obtained by an independent method). Note that the uncertainty of this independently obtained total ozone value is not currently known to the author; however, even if quite erroneous, it serves its purpose in this retrieval simulation as it separates the true and first guess profiles. The average tropical atmosphere O_3 profile was then used for a first guess in the retrieval process. Figure 5.3 illustrates vertical profiles of ozone and temperature for the average tropical atmosphere and the Brazzaville (BZV) sounding.

As discussed in Section 4.5, the O_3 energy contribution functions (O_3 ECFs) indicate the vertical regions from which information content arises in the measured signal; and their broadness near the maxima dictate approximate vertical length scales of resolution about the corresponding peak contributions. By examining the broad O_3 ECFs for this measurement (Figure 4.16), one can see that any hopes for vertical profiling capability are quickly reduced to expectations of slight layering or

simply integrated layer amounts, especially in the troposphere. However, as this retrieval simulation will demonstrate, in vertical regions where the character of the true profile is well represented by the a priori information, one can also expect good agreement between the retrieved and true profile shapes. This is made possible by the number of levels used to specify the O_3 input profile in the forward model ($n = 46$) being large, and actually much greater than simply the small number of levels that one might expect to achieve from viewing the O_3 ECFs; note that this can lead to undesirable oscillations in the retrieved profiles for bad choices of the a priori constraint, but should not significantly degrade the retrieved integrated ozone amounts.

Figure 5.4 summarizes the results of the retrieval simulation for this tropical atmosphere case with elevated tropospheric ozone. After two iterations, the retrieved integrated ozone amounts appear closest to the truth. However, the result of the fifth iteration is defined as the retrieved profile for this simulation, since convergence of signal residuals (i.e., root-sum-square (rss) of the signal residuals) has been used for determining retrieval completion. And as the rss signal residuals show, the outcome does not change with continued iteration since the signals have already been matched to within the measurement uncertainty (equation 5.6) while satisfying the a priori constraint (equation 5.7). As Figure 5.4 demonstrates, the retrieved profile matches the true profile quite well for this particular simulation in both profile character as well as integrated ozone amounts. The inability to match the profile variability in the lower troposphere is an expected consequence of the broad nature of the O_3 ECFs in this vertical region. This simulation serves to exemplify the successful implementation of the maximum likelihood retrieval technique applied to this research and to a certain extent the capabilities of this particular measurement. Note that while

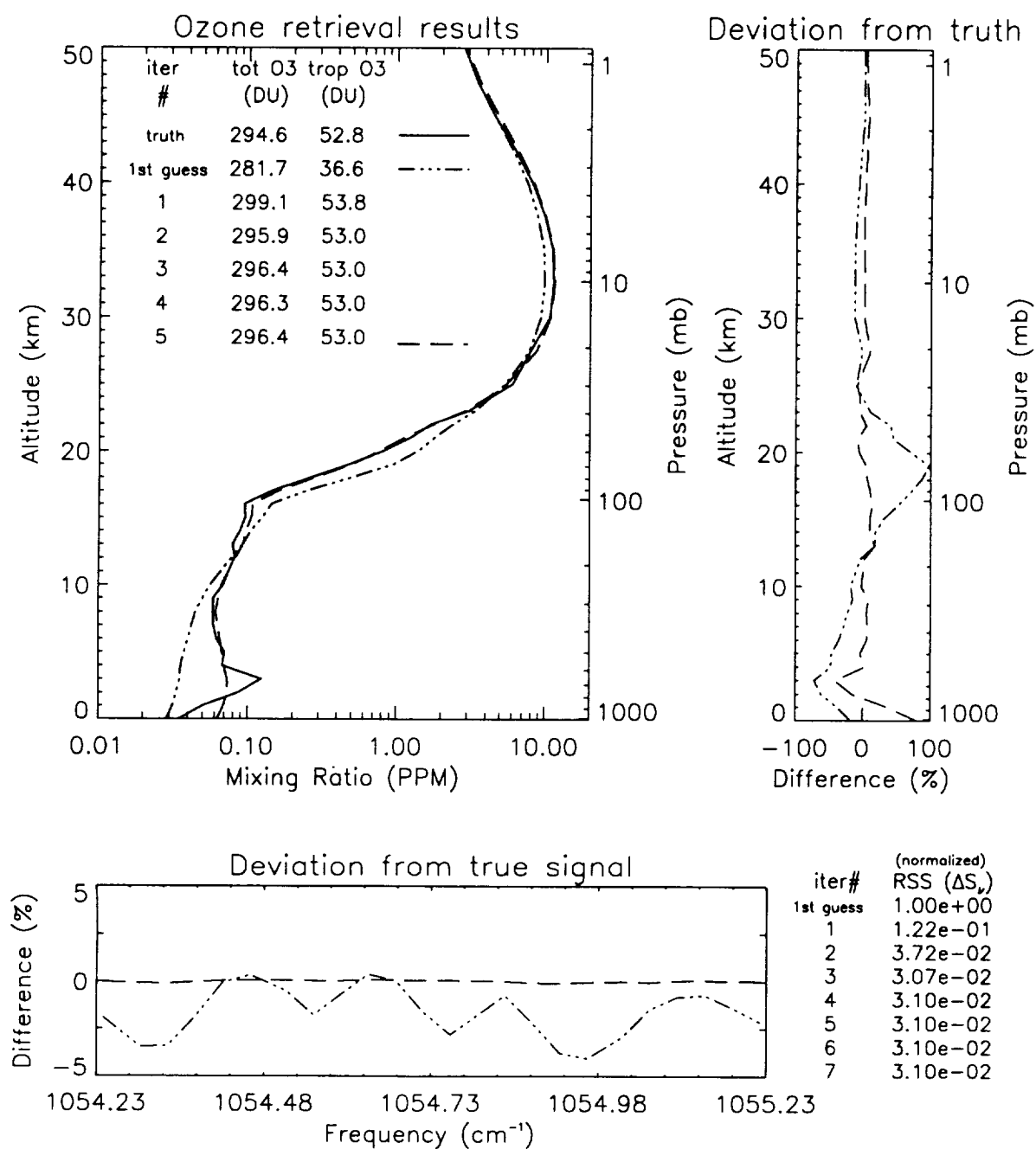


Figure 5.4: Retrieval simulation for a tropical atmosphere having enhanced tropospheric ozone. Sounding data from Brazzaville, Congo, on August 5, 1992, used for true profile; the 1976 Standard Atmosphere (tropical) used for first guess.

these results are very encouraging, they pertain to one particular retrieval and are not representative of “average” expectations. One would expect these results to be better than the average due to the likeness in character between the first guess and true profiles. In addition, while measurement and null-space errors are incorporated into this retrieval, an error analysis must be performed to consider the impact of forward model uncertainties. Section 5.6 will give a more complete assessment of the overall capabilities of this measurement.

5.6 Results of Error Analysis

This section will address the effect of errors on the retrieved state of atmospheric ozone corresponding to measurement with the simulated instrument of this research. The error analysis technique (Section 5.3) will be implemented as discussed in Sections 5.3 and 5.4. The total random error will be estimated by combining the null-space and measurement errors with random components of the forward model error; this can be thought of as the instrument measurement precision. A measurement accuracy will also be estimated by combining the effects of forward model systematic uncertainties.

Accuracy can be defined as closeness to the true value, and it is generally limited by systematic errors; however, it can be thought of as depending on random errors to some extent, since a given accuracy implies a precision at least as good. Precision can be interpreted as how exactly a result is determined (regardless of what the result means), or simply as the reproducibility of a result; the precision is generally limited by random errors. Variability of the retrieved ozone solution will be determined by random independent errors as defined by measurement error, null-space error, and the random part of the forward model uncertainties (i.e., model parameter errors which

can vary randomly from one measurement to the next). The systematic components of the forward model error are constant and do not contribute to variability of the ozone solution vector \mathbf{x} , they just make a non-zero ensemble mean of $\mathbf{x}_{true} - \mathbf{x}_{retrieved}$ which will loosely be considered as accuracy in this analysis.

Although all possible forward model errors have not been simulated, an attempt has been made to include those most significant. The following uncertainties will be considered as random components of the forward model error: surface emissivity $\epsilon_s \pm 1\%$, surface temperature $T_s \pm 1K$, temperature profile $T(p) \pm 2K$ (one level at a time), along with the interfering species $H_2O(p) \pm 30\%$ and $CO_2(p) \pm 5\%$. While the systematic components of the forward model error will be represented by: water vapor continuum model self-broadening coefficient $C_s \pm 10\%$ and nitrogen-broadening coefficient $C_n \pm 10\%$, radiometric signal calibration $\pm 1\%$; sensitivity to uncertainties in spectral line parameters will be assessed with H_2O halfwidths $\pm 5\%$, O_3 halfwidths $\pm 5\%$, H_2O line strengths $\pm 10\%$, and O_3 line strengths $\pm 1\%$. The magnitudes of these uncertainties were chosen based on anticipated knowledge of the forward model parameters, presumably obtained from some independent sources (i.e., laboratory measurements and calibrations, climatological data, independent sensors, etc.). Note that while the \pm signs are shown with the errors, the effect of either $+$ or $-$ each uncertainty will be demonstrated; however, since the actual signs are unknown and may be assumed to vary randomly among the individual errors, total errors will be estimated by taking a root-sum-square (rss) of the component errors (even for those systematic in nature). Some uncertainties included in the random error compilation (i.e., emissivity and temperature of the surface) might also possess systematic errors; their corresponding effects on retrieved ozone will be demonstrated but not included in the overall error estimates, which assume individual errors to be

either random or systematic. In addition, effects from other systematic offsets will be calculated for demonstration purposes. For example, a case showing the impact of completely ignoring the water vapor continuum will be simulated; thus illustrating its importance, especially in the tropical troposphere. Also, the impact of systematic offsets in input profiles (at all levels simultaneously) of temperature and interfering species will be evaluated.

The retrieved ozone sensitivity to each uncertainty will be calculated as a function of altitude, and will then be used in computing the impact on integrated ozone amounts for both the troposphere and total column abundance. In computing the integrated effect, the impact on ozone amount for each layer can be simply summed up for the systematic errors, whereas they may be combined by taking an rss (and thus reducing the impact on integrated amounts) for random errors since the signs may vary independently between errors at different levels (i.e., in the nominal temperature profile).

The annual average tropical atmosphere used as a first guess in the retrieval simulation of Section 5.5 will be the first case considered in this error analysis. We will start by examining some of the random components to the forward model error. Equation 5.14 can be used to estimate their combined effect on the ozone retrieval. In order to accurately determine the sensitivities to random uncertainties in these input parameter profiles, knowledge of their profile covariances (variances at each level as well as correlations among the different levels) is necessary. Note that the impact of random perturbations (of fixed magnitude) is a function of the number of levels at which the input profiles are specified in the forward model. Consequently, perturbation amplitudes deduced from a statistical analysis of typical input profiles would require adjustment if the number of vertical levels requiring specification in

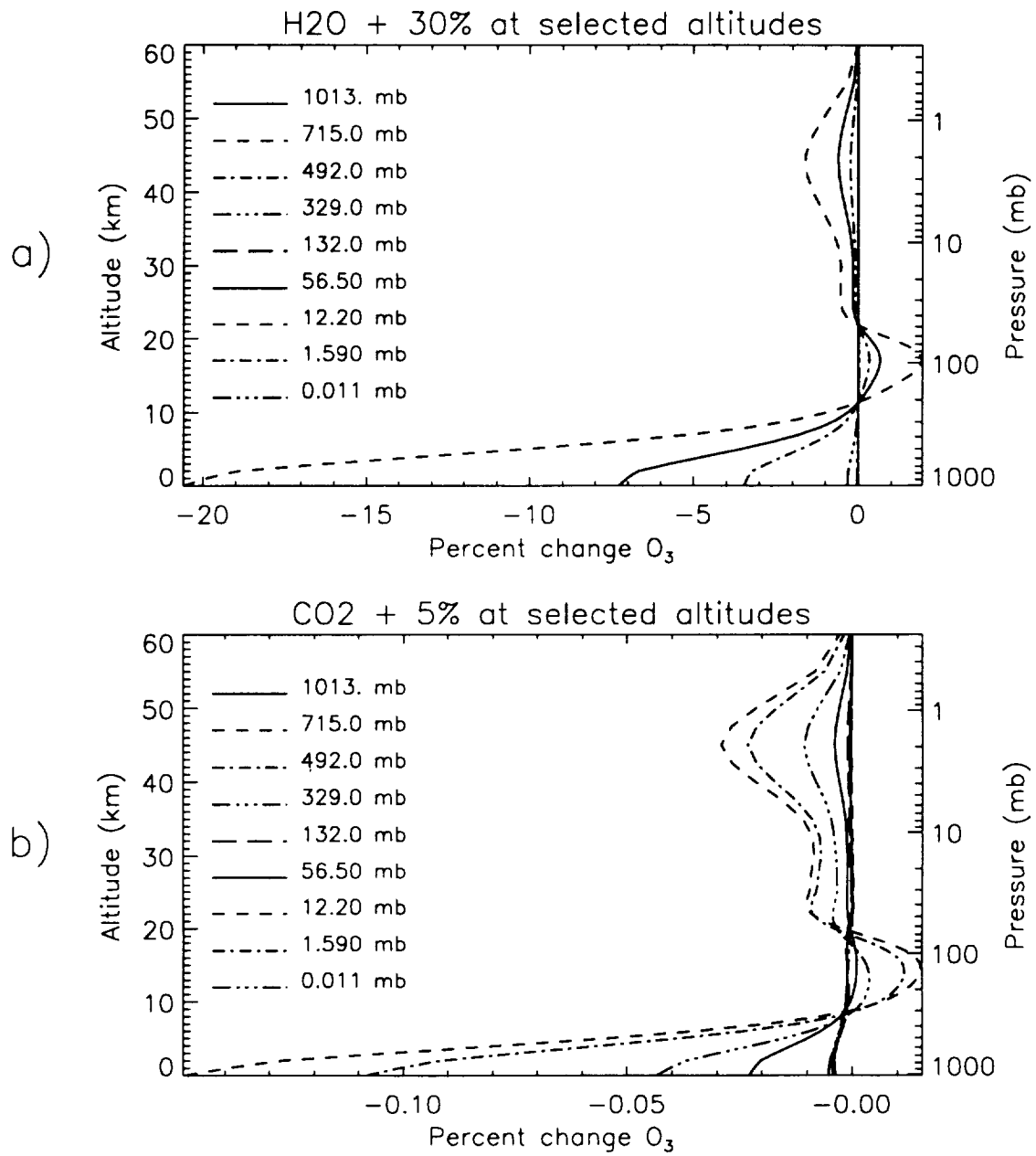


Figure 5.5: Retrieved ozone sensitivity to perturbations of interfering species at specific altitudes for a tropical atmosphere.

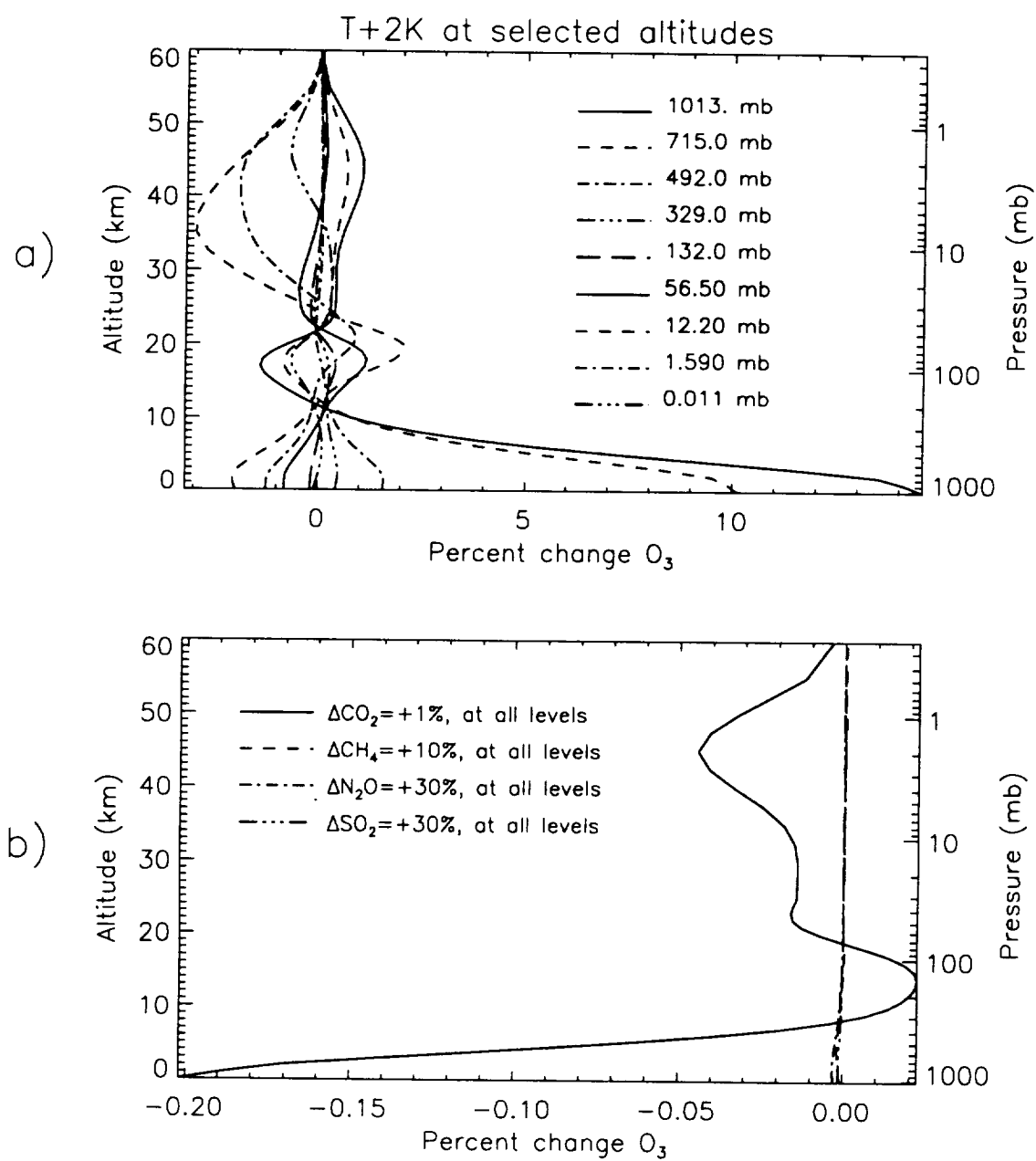


Figure 5.6: Retrieved ozone sensitivity to temperature profile perturbations and systematic offsets in some interfering species for a tropical atmosphere.

the forward model were different from that of the database from which the input profile covariances were derived. For this study, perturbation amplitudes were arbitrarily chosen to represent forward model input profile uncertainties at the various model input levels without including inter-level correlations. Figures 5.5 a), 5.5 b), and 5.6 a) show the retrieval sensitivity to perturbations in knowledge of H_2O , CO_2 , and temperature, respectively, at selected altitudes. Note that a perturbation at a particular altitude also affects the retrieved ozone amount at other altitudes due to the broadness of the O_3 ECFs. In addition, local maxima of perturbation effects exist in the stratosphere and lower troposphere due to peaks in the O_3 ECFs at these altitudes; whereas the local maximum in the upper troposphere region results from this vertical region having low O_3 information content (minima in O_3 ECFs) and high a priori information uncertainty (due to tropopause height variability). The sensitivities to errors in interfering species illustrated in Figure 5.5 show maximum error in O_3 near the surface to result from perturbations slightly above the surface (~ 715 mb). The near-surface sensitivity is greater for perturbations in this above-surface region since they affect the amount of upwelling radiation (reaching the nadir-viewing satellite sensor) originating from this altitude and below; and the sensitivity decreases with further elevation in the troposphere as the temperature and concentrations of interfering species decrease. Note that this decrease happens sooner in the H_2O perturbation case, since the H_2O concentration decreases with increasing altitude more sharply than CO_2 . This feature is not exhibited in the temperature perturbation case because perturbations at the higher temperature (near the surface) have a larger impact, since both the Planck function and transmittance (see equation 4.14) are affected. These plots demonstrate the importance of knowledge pertaining to the H_2O and temperature profiles. The uncertainty in retrieved

Results of Error Analysis: 1) tropical (15N annual average)		
	Effect on retrieved integrated O_3 amounts	
Perturbation	Accuracy (%) (systematic errors)	
	Troposphere	Total
H_2O continuum model		
without H_2O continuum	71.73	12.21
Cs & Cn +10%	-6.485	-1.104
surface emissivity		
ϵ_s -1%	-12.99	-2.187
surface temperature		
T_s +1K	22.32	3.755
temperature profile		
T+1K all levels	10.65	-1.267
T+1K in troposphere & -1K in regions above	16.75	5.531
interfering species		
CO_2 +1% all levels	-0.04610	-0.01887
CH_4 +10% all levels	-0.0008151	-0.0001381
H_2O +10% all levels	-12.21	-2.152
N_2O +30% all levels	-0.001551	-0.0002209
SO_2 +30% all levels	-0.0006869	-0.0001187

Table 5.2: The effect of certain systematic perturbations on retrieved integrated ozone amounts for a tropical atmosphere.

ozone due to a random error in an input profile can be estimated by combining (using an rss) the error profiles resulting from perturbations (of the desired error magnitude) at each altitude of which the forward model input profile is specified. When examining these error profiles, one must keep in mind the broad nature of the O_3 ECFs pertaining to this measurement, which makes the integrated effects (effects on integrated ozone amounts) more meaningful quantities; these integrated effects will be presented in tabulated form later in this section.

Figure 5.6 b) illustrates the effect on retrieved ozone from systematic offsets in some of the interfering species; this shows error contribution associated with lack of knowledge regarding CH_4 , N_2O , SO_2 , and probably CO_2 to be insignificant at

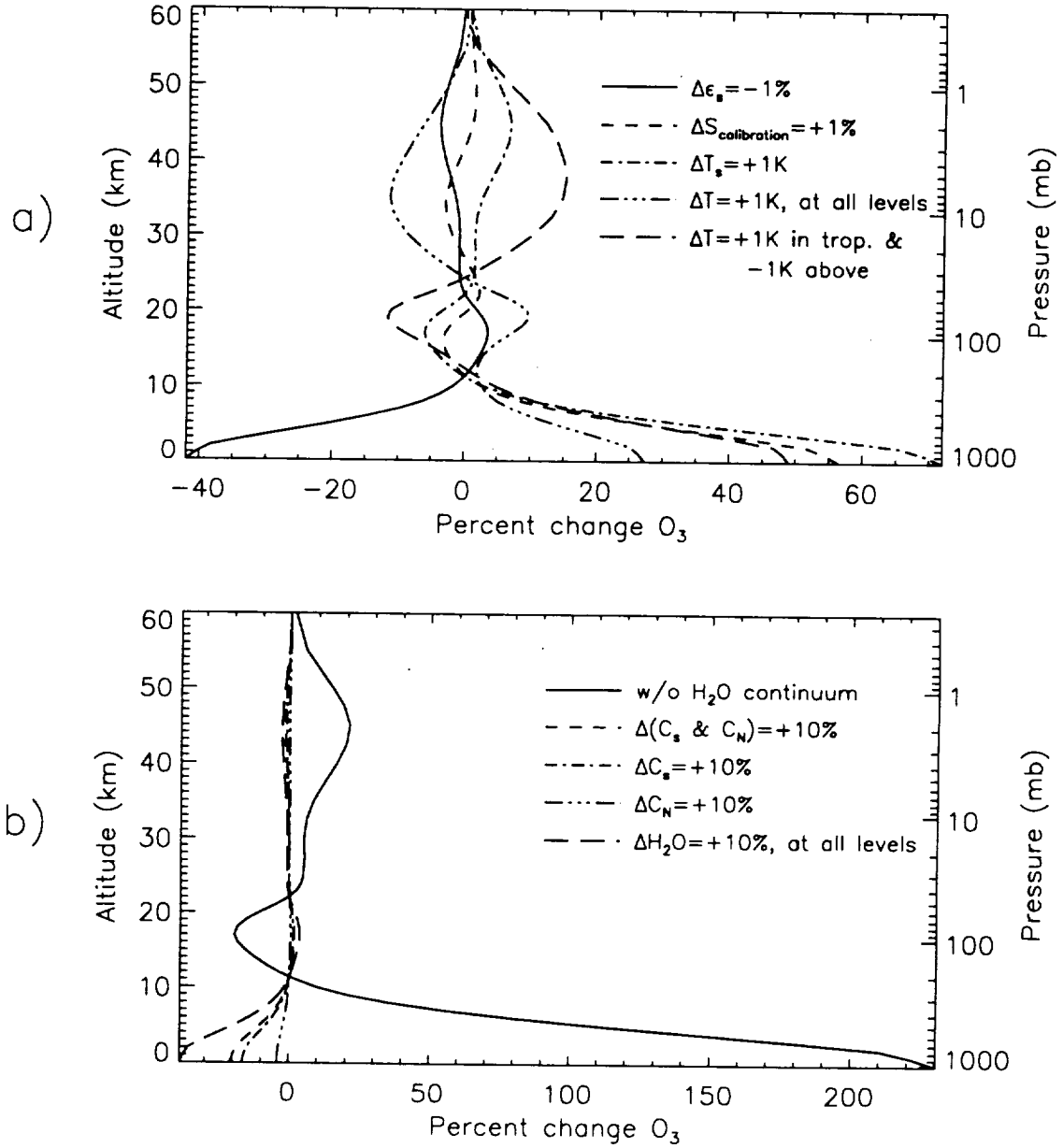


Figure 5.7: Retrieved ozone sensitivity to perturbations in various forward model parameters for a tropical atmosphere.

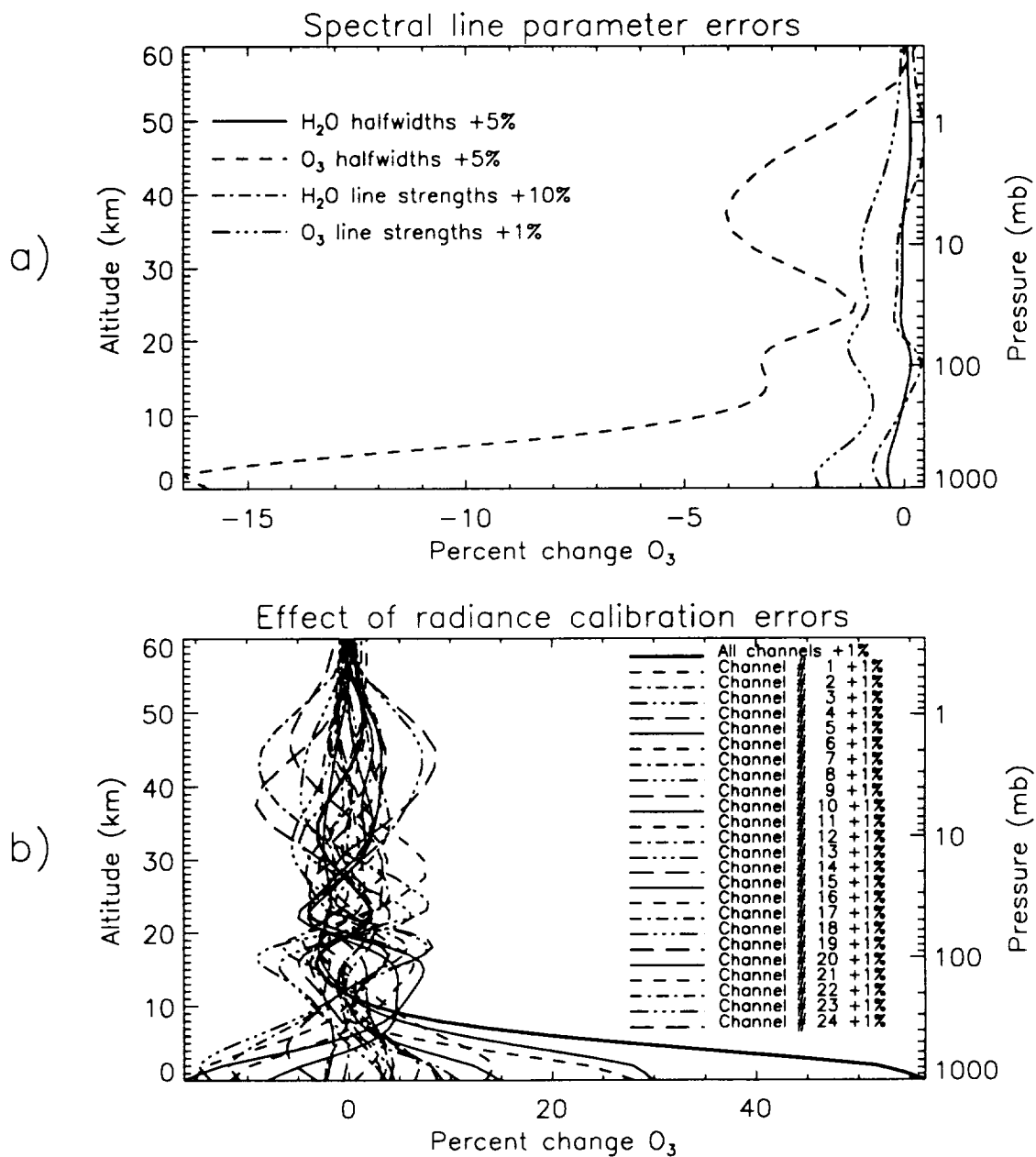


Figure 5.8: Retrieved ozone sensitivity to perturbations in spectral line parameters and radiometric signal calibration errors for a tropical atmosphere.

Results of Error Analysis: 1) tropical (15N annual average)		
Effect on retrieved integrated O_3 amounts		
Perturbation	Accuracy (%) (systematic errors)	
	Troposphere	Total
signal calibration		
all channels +1%	17.97	1.963
channel # 1 +1%	-4.948	2.609
channel # 2 +1%	-4.450	0.7598
channel # 3 +1%	-5.347	0.7472
channel # 4 +1%	-2.030	-0.001413
channel # 5 +1%	4.852	0.2224
channel # 6 +1%	6.412	1.729
channel # 7 +1%	2.034	0.3676
channel # 8 +1%	-0.2036	0.004968
channel # 9 +1%	2.616	-0.3632
channel # 10 +1%	9.612	-0.6286
channel # 11 +1%	6.309	-0.4318
channel # 12 +1%	1.078	-0.03715
channel # 13 +1%	-2.385	1.181
channel # 14 +1%	-0.7961	0.7265
channel # 15 +1%	1.677	-0.2140
channel # 16 +1%	-0.5309	-0.7034
channel # 17 +1%	0.1599	-1.718
channel # 18 +1%	1.337	-2.109
channel # 19 +1%	2.715	-2.252
channel # 20 +1%	-0.1729	-0.6613
channel # 21 +1%	1.117	0.6545
channel # 22 +1%	1.729	0.5754
channel # 23 +1%	-0.4567	0.2314
channel # 24 +1%	-2.361	1.274

Table 5.3: The effect of systematic perturbations of the radiometric signal calibration on retrieved integrated ozone amounts for a tropical atmosphere.

the uncertainty levels simulated, due to avoidance of line overlap in the channel selection process (see Section 4.5). The effects resulting from uncertainties in some other forward model parameters are exhibited in Figure 5.7. Uncertainties in surface emissivity or temperature, such as those shown in Figure 5.7 a), can contribute to systematic or random errors (or both) depending on their source of origination; accordingly, how they contribute to error in integrated O_3 or the method of combining with other errors varies. The radiometric signal calibration error (which has been applied to all channels simultaneously) and the temperature profile offsets would contribute to the systematic error, if present. Note that the case containing both positive and negative temperature biases for the troposphere and stratosphere (and regions above), respectively, corresponds to what is believed to be a worst case temperature error scenario. Such anticorrelated offsets could arise from a temperature retrieval scheme trying to compensate for a temperature error in one particular vertical region by introducing an error of opposite sign in another region, although the regions would probably be of smaller vertical length scale than shown here. The magnitudes of all potential errors in Figure 5.7 a) are significant and, accordingly, effort should be made to minimize chances of their existence. Figure 5.7 b) illustrates retrieval sensitivity to systematic offsets in forward model parameters related to water vapor. The case neglecting the H_2O continuum demonstrates its importance to the radiative transfer within this spectral region for a (moist) tropical atmosphere. However, since this is known, a water vapor continuum model has been incorporated into the forward model; therefore, neglecting it is an unrealistic source of error. Alternatively, the sensitivity to H_2O continuum model coefficients and the actual amount of H_2O must be considered. The self- and nitrogen-broadening coefficients are generally thought of as being good to within 10% (see, for example, Zhao, 1992). Figure 5.7

b) illustrates the impact of such an uncertainty in these coefficients, individually and for both simultaneously. The effect of an offset in the H_2O profile is also shown.

The effect on integrated O_3 amounts of each of the mentioned errors acting individually will be presented, and for a selected subset of these errors a total error impact will be estimated. Table 5.2 shows how some of the previously discussed systematic errors would affect the accuracy, if they were to exist. As already mentioned, the simulation without the water vapor continuum is merely for demonstration purposes. The simultaneous perturbation of the self- and nitrogen-broadening coefficients also fits into this demonstration category, since it is probably more likely that uncertainty in these coefficients would vary independently of each other. Uncertainties in the surface emissivity, surface temperature, temperature profile, and H_2O profile will be assumed random in nature for compiling the total error estimates; however, as represented in Table 5.2, systematic offsets in these forward model parameters can severely impact the accuracy of the retrieved ozone amounts. This table also shows reasonable uncertainties on some other interfering species to be insignificant, as was discussed earlier regarding Figure 5.6 b). Figure 5.8 a) illustrates the retrieved ozone sensitivity to errors in spectral line parameters. In particular, sensitivities to typical uncertainties in halfwidths and line strengths for both O_3 and H_2O are demonstrated. The impacts on retrieved ozone are negative (in sign) since a decrease in O_3 concentration is required to offset the increased absorption which results from each one of these spectral line parameter errors shown. The radiance calibration error shown earlier (Figure 5.7 a)) assuming an error in all channels simultaneously is shown in Figure 5.8 b) along with the impact of the same +1% perturbation acting alone for each of the 24 channels. Note that average pressures from layers which peak energy contributions originate for the instrument channels are contained within Table 4.4.

Table 5.3 summarizes the effect on accuracy of retrieved integrated ozone amounts with these radiometric signal calibration errors included as systematic offsets. The maxima and minima error effects occur at different altitudes for the different channels perturbed separately, since the vertical distribution of O_3 information content varies with the various channels. Considering these radiance calibration errors to act independently would lessen their net error impact, however, a more conservative (and realistic) approach is to assume a systematic offset for all channels simultaneously. Note that the erroneous increase in retrieved tropospheric ozone from the positive perturbation of all signal channels simultaneously results from the algorithm increasing absorption by ozone (of the large surface signal contribution) to counteract the artificial signal increase.

As discussed in Section 5.3, the total random error is composed of the random parts of the forward model error, along with the measurement noise and null-space error. Figure 5.9 illustrates the effects of these main random error components and the total (rss) random error as a function of altitude for this tropical atmosphere case. The measurement noise and null-space error components were computed from equations 5.12 and 5.13, respectively. The random components of the forward model error being considered in this analysis have already been discussed in this section. Figure 5.9 shows the net forward model random error, formulated by taking an rss of its components. This total forward model random error is the most significant random error impacting the troposphere for this tropical case.

The next step of this error analysis is to estimate the impact of these assumed systematic and random errors on the overall accuracy and precision, respectively, for this measurement. Table 5.4 summarizes the component and composite errors on retrieved integrated ozone amounts for the tropical atmosphere case under eval-

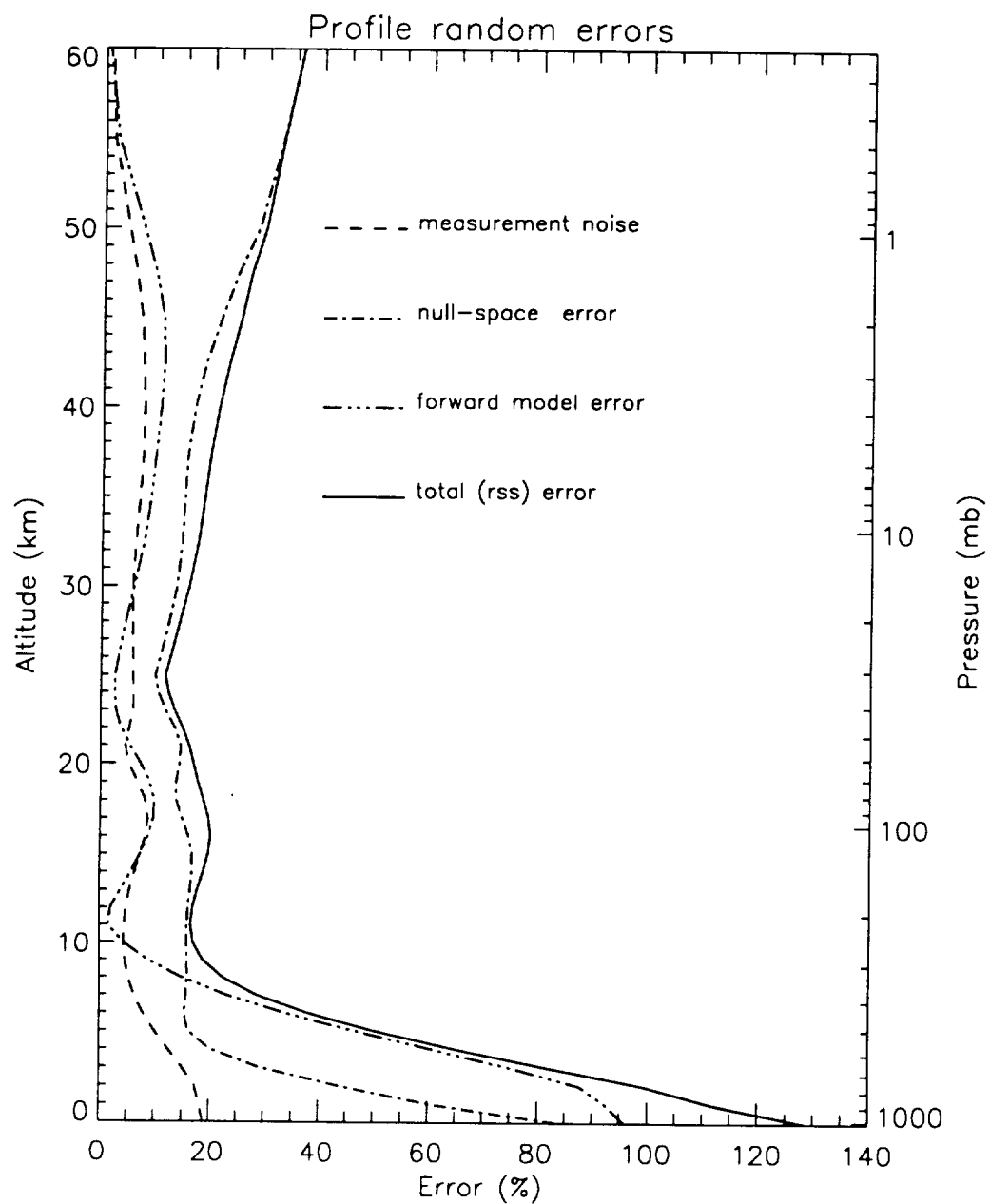


Figure 5.9: Effects of component and total random errors on retrieved ozone for a tropical atmosphere.

Results of Error Analysis: 1) tropical (15N annual average)				
Effect on retrieved integrated O_3 amounts				
Perturbation	Accuracy (%) (systematic errors)		Precision (%) (random errors)	
	Troposphere	Total	Troposphere	Total
H_2O continuum model				
Cs +10%	-5.186	-0.8828		
Cn +10%	-1.314	-0.2239		
surface emissivity $\epsilon_s \pm 1\%$			3.734	0.5333
signal calibration all channels +1%	17.97	1.963		
surface temperature $T_s \pm 1K$			6.415	0.9162
temperature profile $T \pm 2k$ one level at a time			2.574	1.006
interfering species				
$H_2O \pm 30\%$ one level at a time			3.470	0.5087
$CO_2 \pm 5\%$ one level at a time			0.02772	0.005789
spectral line parameters				
H_2O halfwidths +5%	-0.1448	-0.03795		
O_3 halfwidths +5%	-8.379	-3.103		
H_2O line strengths +10%	-0.2331	-0.1155		
O_3 line strengths +1%	-1.245	-0.9663		
Error impact (rss)				
forward model uncertainties	20.61	4.024	8.589	1.548
measurement noise			1.978	1.206
null-space error			5.514	2.648
total (rss) error	20.61	4.024	10.40	3.298

Table 5.4: Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a tropical atmosphere.

uation. Therefore, based upon the uncertainties assumed and the error compilation procedure implemented, one would expect this instrument to enable (tropical) tropospheric ozone determination to within $\sim 10\%$ and $\sim 20\%$ for precision and accuracy, respectively; and knowledge of total column ozone abundance to within $\sim 3\%$ and $\sim 4\%$ for precision and accuracy, respectively. Note that this analysis has been performed for the annual average tropical atmosphere and that tropical atmospheric cases of most interest would probably contain elevated tropospheric ozone, for which the sensitivity to these uncertainties and the overall error in retrieved ozone amounts would be reduced. Also, it may be possible to perform the signal calibration to better than 1% and thus significantly improve the accuracy of this measurement.

This error analysis procedure was repeated for the remaining five atmospheric types introduced in Section 5.4. A brief discussion of these final results will now follow. Summary tables similar to Tables 5.2 and 5.4 have been formulated for each of the additional atmospheres and are included in Appendices B and C, respectively. Upon examining the sensitivities to systematic error perturbations of the various atmospheres presented in Appendix B, one can observe some expected results. For example, the water vapor continuum drastically decreases in importance as the focus goes from a warm, moist tropical atmosphere all the way to a cold, dry subarctic winter atmosphere. Also, the significance of uncertainties in surface emissivity and temperature decreases as one considers atmospheres less sensitive to surface contribution variations (i.e., a winter atmosphere). Appendix C presents the total error budget (as done for the tropical atmosphere) for the other atmospheric cases. Some notable observations are as follows. The sensitivity to signal calibration errors increases for atmospheric cases with larger surface signal contributions (i.e., warm atmospheres), for which a perturbation of total signal becomes a larger part

Results of Error Analysis: total (rss) errors				
Effect on retrieved integrated O_3 amounts				
Atmospheric type	Accuracy (%)		Precision (%)	
	(systematic errors)		(random errors)	
	Troposphere	Total	Troposphere	Total
1) tropical (15N annual average)	20.61	4.024	10.40	3.298
2) mid-latitude summer (45N July)	15.12	3.278	7.730	3.311
3) mid-latitude winter (45N Jan)	11.94	1.987	8.627	3.681
4) subarctic summer (60N July)	15.31	2.932	8.224	3.317
5) subarctic winter (60N Jan)	7.549	2.235	10.32	3.945
6) U.S. standard (1976)	16.38	2.779	9.656	3.421
mid-latitude [2) & 3)] average	13.53	2.633	8.179	3.496
subarctic [4) & 5)] average	11.43	2.583	9.272	3.631
overall [1)→5)] average	14.11	2.891	9.060	3.510

Table 5.5: Summary of total (rss) errors in retrieved integrated O_3 amounts for all atmospheric cases considered.

of the “ozone signal.” The null-space error results from vertical regions that cannot be measured by the observing system, so this error is largest for atmospheric cases having the least variation in thermal structure (i.e., the subarctic winter case). One can also notice that the measurement noise is smallest for the tropical atmosphere, which possesses the highest signal-to-noise ratio.

Table 5.5 summarizes the effect of all assumed errors on retrieved integrated O_3 amounts for each of the various reference atmospheres considered. Approximate an-

nual means have been computed by averaging the July and January results for both the mid-latitude and subarctic cases. In addition, an overall (northern hemisphere) average has been formulated using the results from atmospheric types 1) through 5). The accuracies have been estimated with knowledge of the retrieval sensitivity to forward model systematic uncertainties which, in general, become more sensitive (and thus less accurate) when the atmosphere has more moisture or higher temperatures. The corresponding precisions vary for these atmospheric cases in a less obvious manner due to the different behavior exhibited by its three components of random error. In terms of magnitudes, the author finds these results to be very encouraging; in fact, the tropospheric values are better than originally anticipated. In addition, these predictions compare quite favorably with capabilities of the limited number of instruments/techniques currently enabling similar observations. For example, measurements from the High-Resolution Infrared Sounder (HIRS), which is part of a set of three instruments defined as the TIROS Operational Vertical Sounder (TOVS), can be used to determine total ozone to within $5 \rightarrow 7\%$ and $5 \rightarrow 9\%$ for precision and accuracy, respectively (Ma et al., 1984; Planet et al., 1984); the ozone $9.6 \mu\text{m}$ band is also sensed by HIRS, however, the retrieval process utilizes correlative data from the UV spectral region (i.e., Dobson spectrophotometer measurements). Data from the Total Ozone Mapping Spectrometer (TOMS) have been improving over its years of operation, mainly from improved characterization of instrumental uncertainties and modifications to the data reduction software; accordingly, the current estimates of total ozone can be good to within $\sim 2\%$ and $\sim 5\%$ for precision and absolute accuracy, respectively, depending on solar zenith angle (Bhartia, private communication, 1993). Although tropospheric ozone is not currently measured directly from any satellite-based instrument, tropospheric ozone residuals can

be deduced from a technique utilizing independent measurements from two satellite instruments. Fishman et al. (1990) have used measurements from the Stratosphere Aerosol and Gas Experiment (SAGE) and TOMS instruments to infer tropospheric ozone integrated amounts, and report deriving the climatology of tropospheric ozone to within an accuracy of $\sim 10\%$ in the tropics and $\sim 15\%$ at mid-latitudes; note that these accuracies refer to annual mean values of tropospheric O_3 residuals, and that the monthly means may typically be accurate to within $\sim 30\%$ for this residual technique. As a final comparison comment, it is worth mentioning again that the estimated tropospheric errors associated with this study would be smaller for atmospheres containing more tropospheric ozone; this would be the case for current typical atmospheres (since nominal amounts of tropospheric ozone have increased since formulation of the reference atmosphere database used in this study), as well as for atmospheric scenarios of particular interest possessing enhanced tropospheric ozone.

CHAPTER VI

SUMMARY AND CONCLUSIONS

The ability to monitor the global distribution of tropospheric ozone (O_3) would be beneficial for enhancement of basic scientific understanding as well as to possibly enable minimizing the occurrences of ill-health effects associated with exposure to elevated amounts of O_3 in the lower atmosphere. Such a measurement capability can be achieved using a satellite-based device performing at high spectral resolution with a large signal-to-noise ratio (SNR), to enable observation in the pressure-broadened wings of strong O_3 lines while minimizing the impact of associated undesirable signal contributions introduced from the terrestrial surface, interfering species, clouds, etc. The Fabry-Perot interferometer (FPI) provides high spectral resolution and high throughput capabilities that are fundamental to the success of this measurement task. An FPI optimized for tropospheric O_3 measurements can simultaneously observe a stratospheric component and thus the total O_3 column abundance, through proper spectral design of the instrument channels. Concentrations of stratospheric O_3 are known to be decreasing; this may lead to an increase in biologically harmful solar ultraviolet (UV) radiation reaching the earth's surface, which would cause detrimental health effects.

A review of the necessary theory of Fabry-Perot interferometry has been pre-

sented in this thesis, including characteristics unique to ideal etalons, real etalons, and instrument configurations involving multiple etalons in series. In light of the intended design, emphasis has been placed on a double-etalon series configuration; accordingly, an expression for a double-etalon FPI instrumental transfer function has been formulated, including the associated broadening effects and incoherent reflections between etalons. A conceptual double-etalon instrument design for measuring atmospheric ozone has been developed in this project. A double-etalon configuration in conjunction with an ultra narrow bandpass filter enable single-order passband transmission with high spectral resolution to achieve the desired spectral isolation. The central order full-width at half-maximum (FWHM) (or resolution) is $\sim .054 \text{ cm}^{-1}$ for the etalons alone and an overall resolution of $\sim .068 \text{ cm}^{-1}$ results after inclusion of aperture broadening introduced from the detection process. The new Circle to Line Interferometer Optical (CLIO) system has been incorporated into this design to efficiently handle the IR-FPI detection task while enabling multichannel operation. The CLIO system converts circular fringes of an FPI into a linear pattern which can be detected by a conventional linear array detector. Other significant instrument design specifications include: a bandpass filter FWHM $\approx 5.27 \text{ cm}^{-1}$, defect finesse of ≈ 37.5 for each etalon, plate coating reflectivities of .92 (corresponding to reflectivity finesses of ≈ 37.7), an aperture finesse of ≈ 37.7 ; and free spectral ranges of 1.52 cm^{-1} and 5.46 cm^{-1} for the high-resolution etalon and low-resolution etalon, respectively. Optical element and system transmittances are as follows: a peak transmittance of .65 is assumed for the bandpass filter, the transmission through each defect-broadened etalon is $\sim .73$, .95 transmittance of the medium between etalons (which is so placed for reduction of inter-etalon reflections), and a value of .85 is assumed for the transmittance of the remaining optical elements not consid-

ered; this corresponds to an instrument peak transmittance of $\sim .28$. This design implies an integrated white light amount of .037 in the optical system, and an overall effective system finesse of ~ 22.4 . It is also worth noting that these instrument design specifications are all obtainable within capabilities of current technology.

A method has been presented for calculating the upwelling radiance reaching a satellite sensor; the target gas ozone and all significant interfering species have been included. The water vapor continuum absorption provides the most significant impact of all interfering species (i.e., interfering spectral lines can be avoided in selection of the channel spectral locations and thus minimize their effects) in this $10\text{ }\mu\text{m}$ spectral region; consequently, modeling of its inclusion into the radiative transfer has been discussed in much detail. The desired instrument spectral resolution and approximate placement for the measurement band was determined by results of simulations assessing the radiance sensitivity to atmospheric ozone perturbations. Several criteria were then considered in determining the exact band placement as well as spectral locations for individual channels: ozone energy contribution functions (O_3 ECFs) needed to demonstrate the existence of sufficient information content (while maxima were desired) in both the troposphere and stratosphere, effects of interfering species should be minimized, and a maximum sensitivity to perturbations in atmospheric ozone amount (especially for the troposphere) was desired. This phase of the research resulted in 24 channels being chosen within the 1054.23 to 1055.23 cm^{-1} spectral region.

A signal-to-noise ratio (SNR) analysis has also been performed, which required definition of other design parameters. Assuming a 50 km nadir footprint along with orbital and instrument scan parameters similar to those of the Total Ozone Mapping Spectrometer (TOMS) instrument aboard Nimbus-7 (achieving continuous coverage

at the equator) yields a dwell time of ~ 220 ms for each target scene. A conservative telescope size of 9 inches in diameter has been assumed which then implies etalon plates of diameter ~ 13.3 cm. A design goal of an effective system f-number ≥ 1.2 has been reached. In order to achieve the maximum attainable SNR, the detector detectivity (D-star) values are desired as high as possible. Consequently, this design incorporates mercury cadmium telluride (HgCdTe) photoconductive (PC) detector elements within the linear array with cryogenic operation at 77 K, utilizing liquid nitrogen (LN_2) cooling and 60° FOV cold stop shielding; this enabled a detectivity $D^* = 5.5 \times 10^{10} \text{ cmHz}^{\frac{1}{2}}/W$ to be assumed for each detector element. The resulting SNR values are a function of the thermal characteristics of the atmosphere and surface below (in addition to the spatial distributions of radiatively active gases within the atmosphere); and for the U.S. Standard (1976) atmosphere correspond to a minimum-maximum range of approximately 133.2 to 289.9 for the various channels.

The maximum likelihood method has been implemented for an ozone retrieval technique in this research. The ability to successfully retrieve integrated ozone amounts for both the troposphere and total column abundance, in the presence of measurement noise and null-space error, has been demonstrated for this conceptual measurement. It has also been shown that the retrieved profile shape can match the true profile shape quite well in some situations (i.e., when the character of the true profile is well represented by the a priori information). An error analysis was performed to ascertain how good the retrieved ozone amounts would be in a statistical sense (encompassing many retrievals) while including the effects of measurement, null-space, and forward model errors. This was applied to six different reference atmospheres to acquire a feeling for the latitudinal and seasonal dependence of errors in retrieved ozone. Results show the proposed instrumentation to enable

a good measurement of actual ozone amounts and even a better determination of relative changes, with a range of accuracy to within 7.55 to 20.6% for integrated tropospheric amounts (and 1.99 to 4.02% for total O_3 column abundance) and a corresponding range in precision to within 7.73 to 10.4% (and 3.30 to 3.95% for total O_3 column abundance), for various atmospheric types considered. Note that these are predicted ozone errors based upon assumed uncertainties in the observing system (measurement process and retrieval technique) and consequently may vary slightly. Nevertheless, these predictions compare quite favorably with capabilities of instruments currently providing total ozone observations or techniques enabling deduction of integrated tropospheric ozone amounts. It is also worth mentioning that the estimated tropospheric errors associated with this study would be smaller for atmospheres containing more tropospheric ozone; this would be the case for current typical atmospheres (since nominal amounts of tropospheric ozone have increased since formulation of the reference atmosphere database used in this study), as well as for atmospheric scenarios of particular interest possessing enhanced tropospheric ozone. Other significant findings include: the importance of including a water vapor continuum model, especially for a warm, moist tropical atmosphere; the significance of uncertainties in surface contribution characteristics (i.e., emissivity and temperature); and great care must be taken in performing the radiometric signal calibration, as retrieved ozone amounts are quite sensitive to such errors.

In conclusion, this thesis project has demonstrated that the Fabry-Perot interferometer used in a double-etalon series configuration in conjunction with the new Circle to Line Interferometer Optical (CLIO) system is a feasible new approach for a satellite-based atmospheric sensor of tropospheric and total ozone. The FPI has a proven success record in previous space applications. While the CLIO subsystem

has only been demonstrated in the laboratory thus far, it is a desirable satellite instrument component as it enables “spatial scanning” through multichannel operation without moving parts. In addition, all suggested design parameters are achievable based on current technological capabilities or expectations, making the proposed instrumentation even more attractive.

As with any feasibility study, additional effort is still needed toward this particular application. From a science perspective, it would be beneficial to determine the effect on retrieved ozone amounts of uncertainties in other forward model parameters, even for those model parameters which have not been included in the current study. For example, the impact of various cloud types and their spatial distributions when cloud cover is present within the instrument field-of-view (FOV), the sensitivity to variations in bandpass filter spectral positioning and shape, the sensitivity to variations in surface topography, and the impact of atmospheric aerosols and chlorofluorocarbons (CFC's). In addition, it would be interesting to assess this instrument's capabilities when applied to a depleted polar stratospheric ozone scenario. Regarding engineering aspects, continued laboratory studies of the CLIO system are recommended, with particular emphasis toward IR applications. It would be helpful to explore pre-spaceflight etalon mounting techniques which minimize surface distortions and thus avoid defect finesse degradation. In addition, it is suggested that ray-trace simulations be performed using optical design software to characterize exact energy distributions at the detector plane; this would enable a more accurate determination of detector element widths and minimum necessary acceptance angles for designing the detector cold stop FOV's, both of which would improve the SNR estimates.

APPENDICES

APPENDIX A

THEORY OF FABRY-PEROT INTERFEROMETRY

A.1 Introduction

Although early work done by George Airy with interference of radiation can be traced back to 1831, the birth of the Fabry-Perot interferometer is attributed to articles by Charles Fabry and Alfred Perot in 1897 and 1899 (Vaughan, 1989). The multiple beam Fabry-Perot interferometer was developed primarily as an improvement on the two-beam Michelson interferometer. The two-beam interferometer produces fringes which are relatively diffuse whereas the multiple beam instrument yields much sharper fringes; this was a great advantage when the eye was being used as a detector (Chantry, 1982). The Fabry-Perot interferometer (FPI) is one of the most flexible and versatile of the multiple-beam interferential devices. Because of the large luminosity-resolving power product (LRP), it has been the instrument of choice for many high-resolution spectroscopic investigations, where the sources are usually intrinsically weak. It is commonly used as a high-resolution spectrometer, a resonant cavity in laser physics, a tuning selector in very high-resolution spectroscopy, etc., for a wide range of applications. Depending upon the application, various versions of the same instrument are implemented; accordingly, the corresponding name may

vary but is usually referred to as one of the following: Fabry-Perot interferometer, Fabry-Perot etalon, tunable etalon, confocal etalon, solid etalon, Piezoelectric Transducer (PZT) scanning resonant cavity interferometer, fixed air gap etalon, or pressure tuned Fabry-Perot. However, this instrument is most commonly known as simply a Fabry-Perot interferometer. The FPI is an extremely powerful tool that has been employed in various forms for different applications over a wide range of wavelengths during the last several decades. As shown by Jacquinot (1954), the principle advantage of the FPI is that its throughput is much larger than that of a prism or grating instrument having the same spectral resolution. This appendix contains an overview of Fabry-Perot interferometry concepts essential to this research project, beginning with characteristics of an ideal FPI (one with perfectly flat, parallel plates) and then discussing the impact of accounting for realistic plate imperfections.

A.2 Characteristics of an Ideal FPI

A Fabry-Perot interferometer is constructed with two partially transmitting mirrors which are parallel to each other and may be flat (plano FPI) or curved (spherical or confocal FPI); this forms the FPI cavity. The “radiused” mirrors in a spherical FPI are separated by their radius of curvature for the special case of a confocal FPI. In a solid etalon, the two surfaces of one substrate are highly parallel and have partially transmitting coatings. The plane Fabry-Perot is comprised of two parallel flat, transparent plates coated with thin films of high reflectivity and low absorption, and is frequently called a Fabry-Perot etalon when the separation between the plates is fixed. An etalon has two mirrors, or plates as they are frequently called, which have partially transmitting coatings on their first surfaces and anti-reflection coatings on their second surfaces. The plates are arranged with their high reflectivity

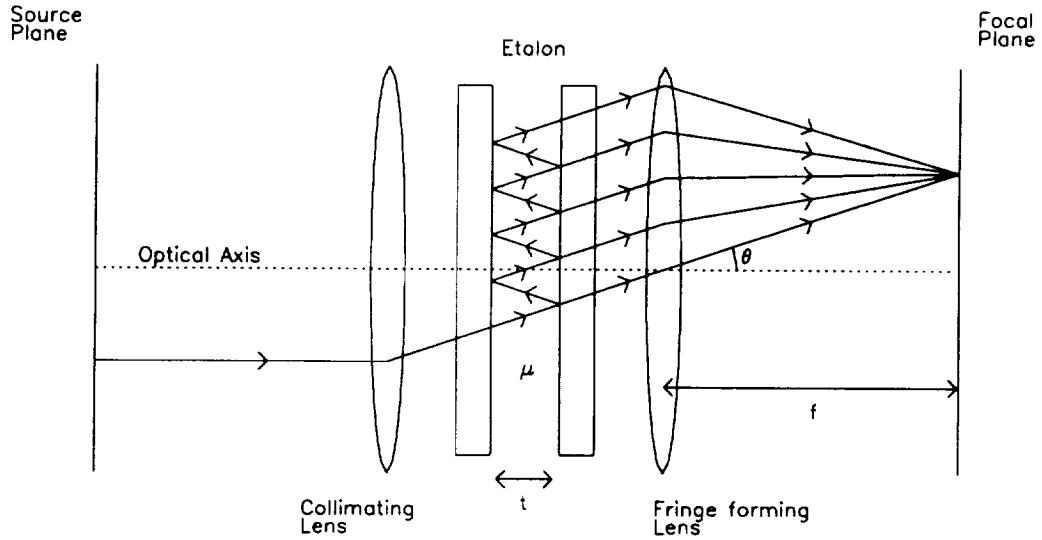


Figure A.1: Schematic diagram of monochromatic ray propagation through a simple Fabry-Perot interferometer system. Light from the source is collected by the collimating lens. The energy undergoes successive reflections within the etalon cavity. The multiple waves are then superposed by the fringe forming lens which forms fringes in the focal plane.

first surfaces facing each other. In the plano (or plane) FPI, the second surface is slightly wedged with respect to the first surface with a wedge angle $\approx 30'$ (Vaughan, 1989) to avoid forming additional (resonant) cavities due to reflections from the low reflectivity surfaces. The plane Fabry-Perot is the simplest and most versatile of the FPI instruments. Etalons of this kind will be the type of FPI's simulated in this research project.

The space between the coated surfaces of the etalon forms a cavity which is resonant at certain wavelengths dictated by the optical properties (i.e., thickness t and refractive index μ) of the gap as well as the incidence angle θ of radiation within the cavity. Due to the low transmission coating, only a small portion of the incident

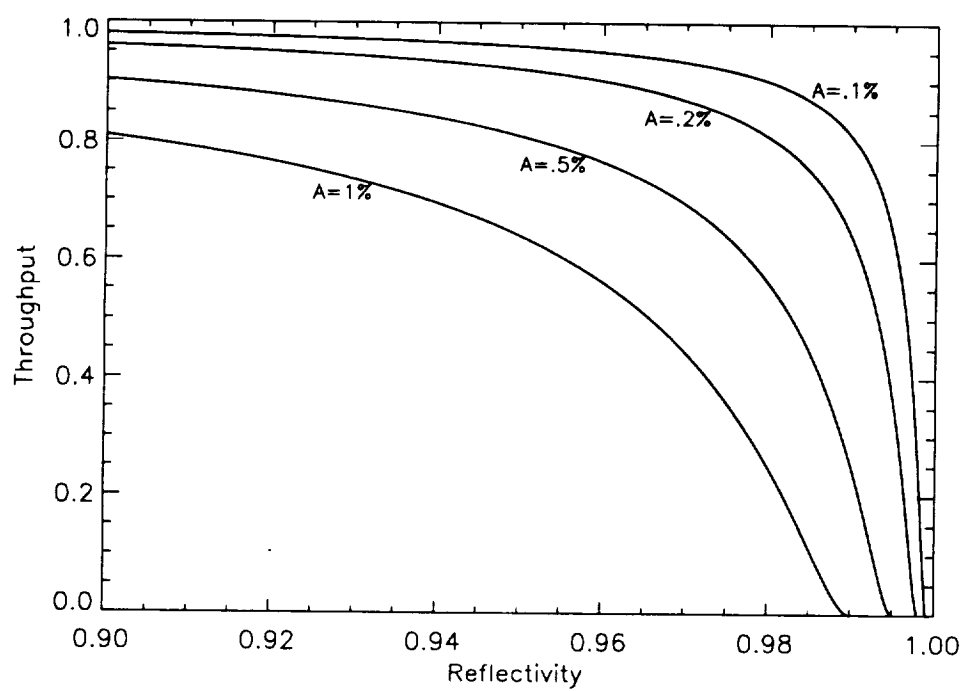


Figure A.2: Ideal FPI throughput versus reflectivity for several values of plate absorption.

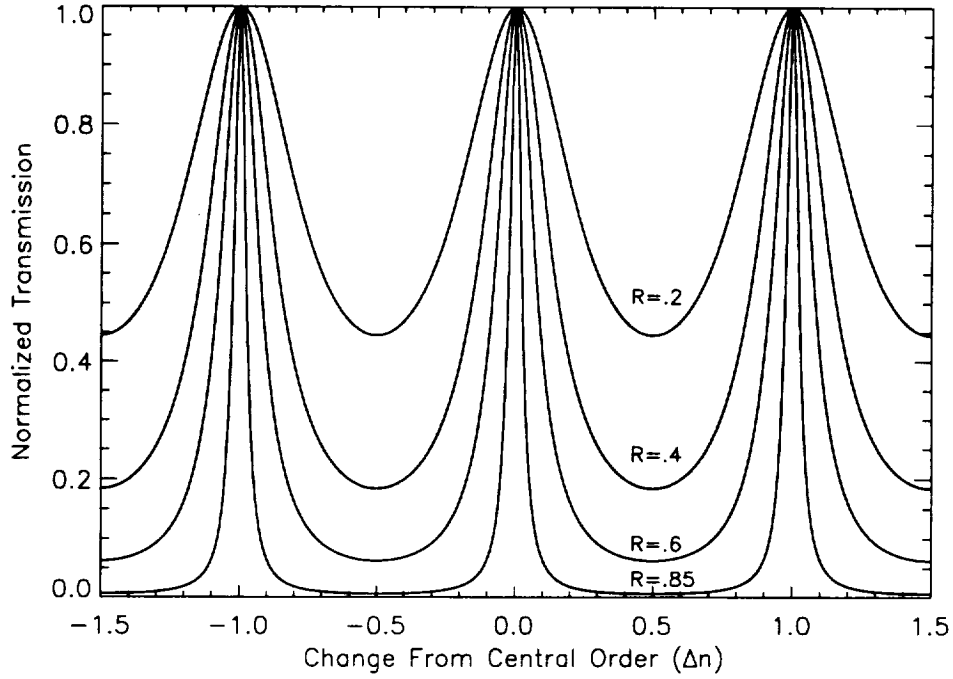


Figure A.3: Transmission through ideal etalon for different plate reflectivities.

radiation enters the cavity where it then undergoes multiple reflections. A simplified schematic of ray propagation through an FPI system is shown in Figure A.1. The amplitude and phase of radiation leaving the cavity depend upon wavelength. Constructive interference produces transmission maxima at resonant wavelengths; conversely, at wavelengths in between these maxima, destructive interference causes most of the incident energy to be reflected thus producing transmission minima. The general equation for transmission through an ideal FPI (assuming parallel, perfectly flat plates with no surface defects; thus having broadening caused only from reflectivity of the surface coatings) is given by the Airy formula for transmission (Born and Wolf, 1975):

$$\mathcal{A}(\Psi) = \tau_a \frac{(1 - R)^2}{1 - 2R \cos \Psi + R^2}, \quad (\text{A.1})$$

where τ_a is a transmission (or maximum throughput) coefficient that includes the

effects of absorption and/or scattering by the reflective coatings, R is the plate reflectivity, $\Psi = 4\pi\mu t\nu \cos \theta + 2\phi$, μ is the index of refraction of the medium between the plates, t is the spacing (or gap) between the plates, ϕ is the phase change upon internal reflection, ν is the wavenumber of incident radiation, and θ is the angle of incidence of the incident radiation. The transmission coefficient τ_a corresponds to the throughput of the ideal FPI at resonance (or its peak transmission) and for an etalon of plate reflectivity R is given by

$$\tau_a = \left(1 - \frac{A}{1 - R}\right)^2, \quad (\text{A.2})$$

where A corresponds to losses due to scatter and absorption for a single plate. The quantity $\frac{\mathcal{A}(\Psi)}{\tau_a}$ (equation A.1 normalized to peak transmission) is defined as the Airy shape function or simply “Airy function,” however, $\mathcal{A}(\Psi)$ alone is commonly referred to as the Airy function as well. Figure A.2 shows the dependence of throughput on reflectivity for several values of A ; A is typically $\leq .2\%$ /surface for modern coatings (Burleigh). It can be seen from equation A.1 that the phase changes upon reflection have no effect on the shape of the Airy transmission function distribution, but serve merely to offset the fringes; and for higher orders of interference (≥ 100) they can be ignored for most practical purposes (Vaughan, 1989). Since the phase contributions can be safely ignored for this application, equation A.1 implies that the FPI has a “comb” of transmission resonances with the peaks in transmission occurring in narrow spectral regions at each of a series of wavenumbers given by:

$$\Psi = 4\pi\mu t\nu \cos \theta = 2\pi n, \quad (\text{A.3})$$

where n is an integer known as the order of interference. It is obvious from equations A.1 and A.3 that if the Airy function is plotted as a function of any one of the quantities μ , t , ν , or $\cos \theta$, one would obtain a regular series of peaks occurring wher-

ever $2\mu t\nu \cos \theta = n$, the integer order. This is illustrated in Figure A.3 for several values of the plate reflectivity.

Taking the total derivative of equation A.3, an expression for the local change in order Δn can be derived giving

$$\Delta n = n - n_0 = 2\mu_0\nu_0 t_0 \cos \theta_0 \left(\frac{\Delta \nu}{\nu_0} + \frac{\Delta t}{t_0} + \frac{\Delta \mu}{\mu_0} - \theta_0 \Delta \theta \right), \quad (\text{A.4})$$

where the zero subscript denotes some nearby reference state. For a fixed gap etalon with a fixed (or small) angle of incidence (i.e., for a particular “channel”), equation A.4 reduces to

$$\Delta n = 2\mu_0 t_0 \Delta \nu \cos \theta_0, \quad (\text{A.5})$$

where $\Delta \nu = \nu - \nu_0$.

In discussing an axial-fringe FPI, the Airy function is commonly considered as a function of ν only, with the other parameters fixed. In this case, the spectral spacing between consecutive passbands at a fixed angle of incidence θ , the inter-order distance, is known as the free spectral range ν_{fsr} (FSR). The free spectral range can be obtained from equation A.5 by solving for the wavenumber difference $\Delta \nu$ induced from consecutive integer values of the order n ; it is given by:

$$\nu_{fsr} = \frac{1}{2\mu t \cos \theta}. \quad (\text{A.6})$$

However, for most axial-fringe instrument applications $\cos \theta$ is approximately equal to 1 and the FSR may be accurately represented by:

$$\nu_{fsr} \approx \frac{1}{2\mu t}; \quad (\text{A.7})$$

this is commonly referred to as the “nominal” free spectral range.

The half-intensity width δ_F , which is sometimes loosely called the resolving limit, is the full width at half-maximum (FWHM) of a transmission peak in the FPI instru-

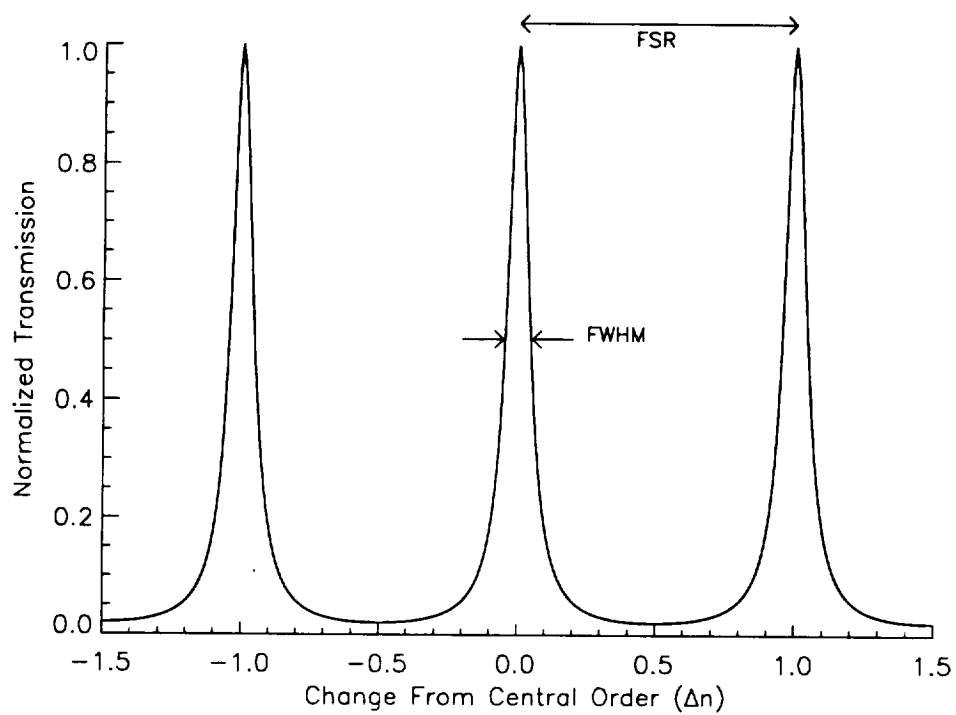


Figure A.4: Transmission through an ideal etalon with the corresponding FWHM (δ_F) and FSR (ν_{fsr}) depicted.

mental function. For an ideal etalon having perfectly flat plates, with only reflectivity broadening existing, this is commonly expressed (i.e., Vaughan, 1989) as:

$$\delta_{Fideal} = \frac{(1 - R)\nu_{fsr}}{\pi\sqrt{R}}. \quad (A.8)$$

The FWHM and FSR of an Airy transmission function are illustrated in Figure A.4. The theoretical spectral resolution, or resolving power, of an FPI is defined by

$$\mathcal{R} \equiv \frac{\nu}{\delta_F}. \quad (A.9)$$

The finesse N is the key measure of the interferometer's spectral resolving capability. The finesse is defined as the ratio of the FSR near the resonance frequency to the FWHM of an etalon transmission peak, and is given by

$$N = \frac{\nu_{fsr}}{\delta_F}. \quad (A.10)$$

Thus higher finesse implies higher resolution, for a fixed FSR. The finesse can also be thought of as the number of passband widths over which the FPI can be scanned (or tuned) without overlapping of orders. For an ideal etalon, N is a function of only the reflectance of the plate coatings and is referred to as the “reflectivity finesse,” N_R , which from equations A.8 and A.10 can be directly represented by

$$N_R = \frac{\pi\sqrt{R}}{(1 - R)}. \quad (A.11)$$

The dependence of reflectivity finesse on plate reflectance is illustrated in Figure A.5.

Another useful parameter is the contrast C which is defined as the ratio of the maximum to minimum FPI transmittance. For an ideal etalon with perfectly flat plates, the contrast is simply the ratio of maximum to minimum Airy function transmission, and can be easily shown to be given by

$$C = \frac{(1 + R)^2}{(1 - R)^2}. \quad (A.12)$$

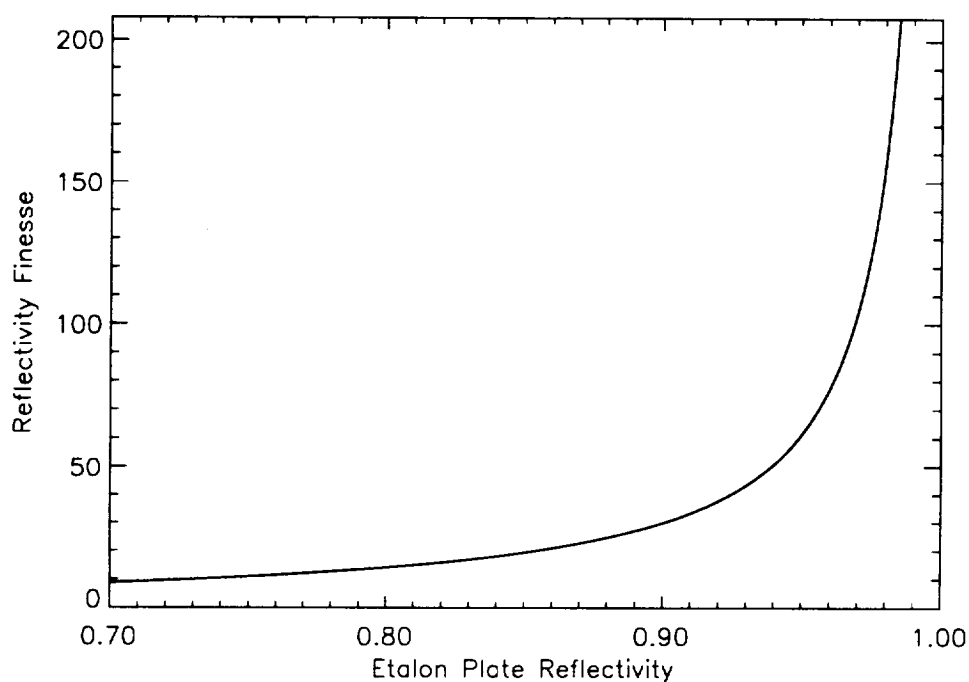


Figure A.5: Reflectivity finesse versus reflectivity for an ideal etalon.

A.3 The Real Etalon

The theory presented in the last section for an ideal etalon assumes parallel, perfectly flat plates. In reality, the Fabry-Perot plates are neither flat nor perfectly parallel; and any deviation in either one (or both) of these quantities will broaden the interferometer passband, requiring modification to some of the equations developed in Section A.2.

Departures from flatness of the interferometer plates, errors in their parallel alignment, and nonuniformities in the reflective coatings cause the effective plate separation t to vary across the surface of the etalon; thus making the optical pathlength vary over the aperture of the interferometer. Consequently, the Airy function (equation A.1) does not coincide for all surface elements within the aperture, and the

resulting fringes are broadened. Broader fringes imply that the same energy must be distributed over a larger area in the focal plane, thus reducing the peak intensity. Therefore, the peak transmission or throughput of a “real” etalon is less than that given for an “ideal” etalon in equation A.2.

The transmission profile of a broadened fringe is necessary in order to simulate a real etalon. This can be determined with knowledge of two functions: the Airy function $\mathcal{A}(\Psi)$ (equation A.1) representing the fringe profile for an ideal etalon and the surface defect distribution function $D(\varphi)$. In these functions, Ψ is the “phase advance” of \mathcal{A} defined by equation A.1, and φ corresponds to the phase decoupling or a perturbation in Ψ . For random surface defects, $D(\varphi)$ can be considered as a probability density function describing the distribution of phase advance Ψ errors φ . For surface defects of a nonrandom nature, $D(\varphi)d\varphi$ corresponds to the proportion of the FPI aperture for which the phase advance differs from some value of Ψ by the amount φ due to plate defects. Therefore $D(\varphi)$ has the attributes of a probability density function with the area under the defect function curve being normalized to unity, denoted mathematically by

$$\int_{-\infty}^{\infty} D(\varphi)d\varphi \equiv 1. \quad (\text{A.13})$$

Accordingly, the mean μ_D and variance σ_D^2 of the defect function can be defined, respectively, by

$$\mu_D = \int_{-\infty}^{\infty} D(\varphi)\varphi d\varphi, \quad (\text{A.14})$$

$$\sigma_D^2 = \int_{-\infty}^{\infty} D(\varphi)(\varphi^2 - \mu_D^2)d\varphi. \quad (\text{A.15})$$

Following an approach similar to Sloggett (1984) and Wilksch (1985), the fractional contribution to the total transmission of the etalon due to those parts of the FPI aperture having phase defect φ can be given by $\mathcal{A}(\Psi - \varphi)D(\varphi)d\varphi$. The total etalon

transmission at Ψ can be obtained upon integration, yielding

$$E(\Psi) = \int_{-\infty}^{\infty} \mathcal{A}(\Psi - \varphi) D(\varphi) d\varphi; \quad (\text{A.16})$$

$E(\Psi)$ can be called the etalon transmission function (ETF). Equation A.16 can be more conveniently represented by

$$E(\Psi) = \mathcal{A}(\Psi) * D(\Psi), \quad (\text{A.17})$$

where $*$ denotes the mathematical process of convolution (see, for example, Churchill, 1972).

The Airy function $\mathcal{A}(\Psi)$ describes symmetrical fringes. If the mean (or central maximum) occurs at $\mu_{\mathcal{A}}$, then from a standard property of convolutions the mean of the defect broadened fringes μ_E can be expressed (Sloggett, 1984) as $\mu_E = \mu_{\mathcal{A}} + \mu_D$. Therefore, the mean of the defect function μ_D is the amount of displacement of the Airy function peaks found in the broadened fringe profile $E(\Psi)$, due to the plate defects. In addition, the peaks (or fringes) described by $E(\Psi)$ will retain their symmetry as long as the defect distribution function is symmetrical about its mean.

Analytical expressions for defect functions representing many common plate imperfections have had many appearances in the literature [e.g., Del Piano and Quesada (1965), Hernandez (1966, 1986), Sloggett (1984), Wilksch (1985), and Vaughan (1989)]. In general, the defect functions are developed by assuming that all imperfections are on one of the reflecting plates which is positioned parallel to the surface of a flawless second plate. Figure A.6 illustrates some of the more typically observed plate imperfections along with their corresponding defect distribution functions most commonly used to represent them. Figure A.6(a) is a situation of flat plates with random defects of small scale and can correspond to plates with poor coatings. The Gaussian distribution function can be used to represent the defect function as this

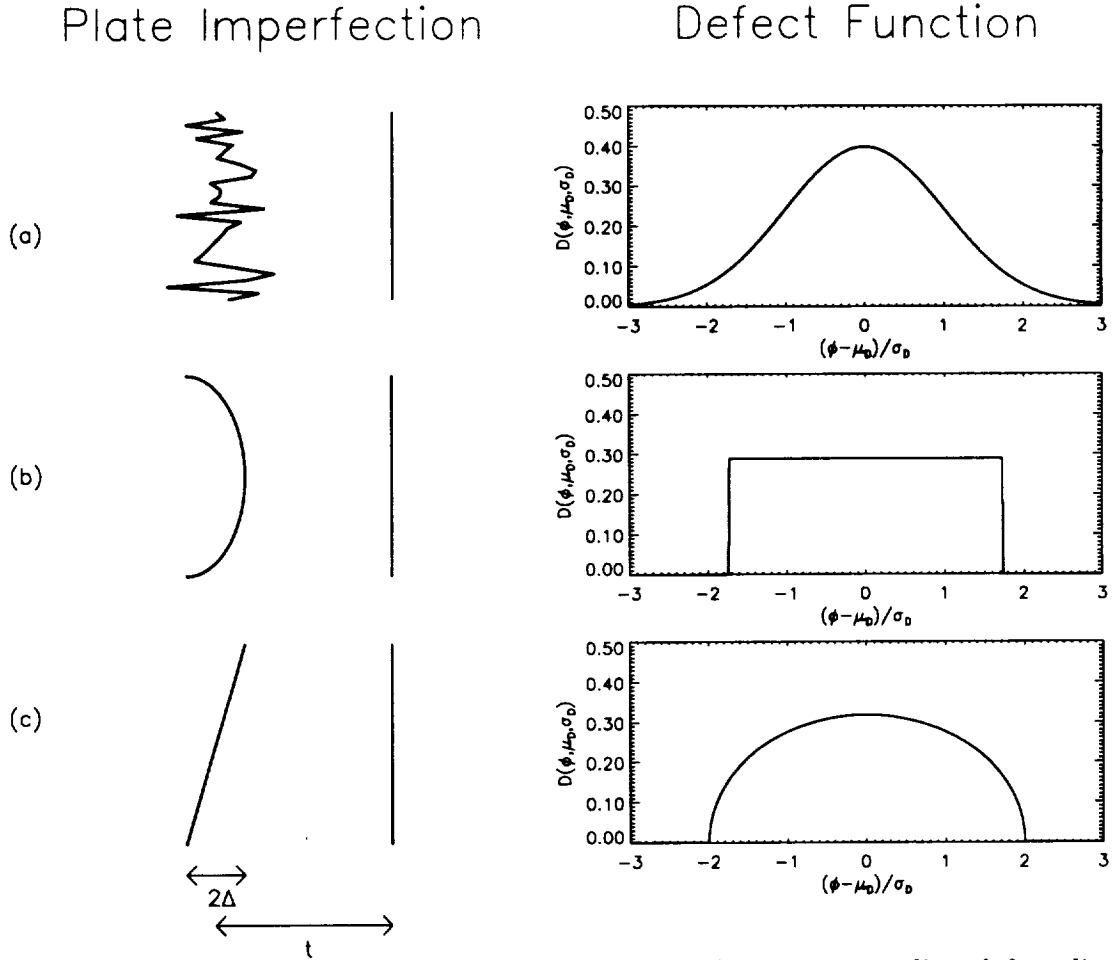


Figure A.6: Common etalon plate imperfections and their corresponding defect distribution functions plotted as a function of phase error normalized to the standard deviation σ_D for each distribution. The spacing between plates (or gap) has a mean of t and a range $t \pm \Delta$ (except points greater than one σ_D from t in the Gaussian distribution). (a) shows a random surface irregularity which can be represented by the Gaussian distribution function. (b) is a case where the plates have a spherical defect and a rectangular function can be used for the defect function. (c) illustrates a pair of flat plates slightly out of alignment, for which a semielliptical function can represent.

Distribution Type	Analytical Expression $D(\varphi)$	Standard Deviation σ_D
Gaussian	$(\sqrt{2\pi}\sigma)^{-1} \exp\left[-\frac{(\varphi-\mu_D)^2}{2\sigma^2}\right]$	σ
Rectangular	$(2\sqrt{3}\sigma)^{-1}$	$\frac{\varphi_{max}}{\sqrt{3}}$
Semielliptical	$\frac{1}{\pi\sigma} \sqrt{1 - \frac{(\varphi-\mu_D)^2}{4\sigma^2}}$	$\frac{\varphi_{max}}{2}$

Table A.1: Analytical expressions used to represents defect distribution functions for some common etalon plate defects. Note that all distribution functions except for the Gaussian vanish for $|\varphi - \mu_D| > \varphi_{max}$, where φ_{max} is the value of φ corresponding to the limit of defect function nonzero extent.

case is that of a small scale random surface irregularity. Figure A.6(b) shows plates with a spherical defect. This type of defect might arise from plates usually being of slightly poorer quality toward the edges compared with the center; in addition, the resultant strain from faulty mounting of the plates can induce this type of curvature effect. For circular plates (or a circular aperture), a rectangular function can be used to represent the corresponding defect function. Figure A.6(c) illustrates a pair of flat plates slightly out of alignment. A semielliptical function can represent its defect distribution for circular plates; and a hemispherical distribution may also be used as a good approximation (Vaughan, 1989). Table A.1 summarizes the above described surface defect distribution functions which are illustrated in Figure A.6. The expressions given for the standard deviations can be obtained through evaluation of equation A.15.

The magnitude of a plate defect is commonly expressed as a fraction of wavelength. This usually corresponds to a maximum excursion from a perfectly plane, parallel surface, and is denoted by 2Δ in Figure A.6. For situations with random

Defect Type	Defect Magnitude	Defect Finesse	Distribution Function Spectral Extent
Random surface imperfections	$(\overline{\delta t_G^2})^{\frac{1}{2}} = \frac{\lambda}{K_G}$	$N_{D_G} = \frac{K_G}{4.7}$	$\sigma_D = \frac{\nu_{far}}{2\sqrt{2} \ln 2 N_{D_G}}$
Spherical curvature	$2\Delta_S = \frac{\lambda}{K_S}$	$N_{D_S} = \frac{K_S}{2}$	$\varphi_{max} = \frac{\nu_{far}}{2N_{D_S}}$
Parallelism error	$2\Delta_P = \frac{\lambda}{K_P}$	$N_{D_P} = \frac{K_P}{\sqrt{3}}$	$\varphi_{max} = \frac{\nu_{far}}{\sqrt{3}N_{D_P}}$

Table A.2: Plate defect characteristics.

surface irregularities, a root-mean-square (rms) deviation $(\overline{\delta t_G^2})^{\frac{1}{2}}$ is normally given to denote microscopic flatness imperfections. A spherical curvature of sagitta is often used as a measure of macroscopic deviations from flatness. The broadening effect of defects on the FPI transmission function can be quantified by a defect finesse for each defect. And the net finesse is obtained by treating the component finessses as parallel impedances. An overall finesse for the above described defects occurring simultaneously can be given by

$$\frac{1}{N_D^2} = \frac{1}{N_{D_G}^2} + \frac{1}{N_{D_S}^2} + \frac{1}{N_{D_P}^2}. \quad (\text{A.18})$$

The spectral extent of the defect distribution function $D(\varphi)$ can be characterized by a standard deviation σ_D and/or φ_{max} , where φ_{max} is the value of φ corresponding to the limit of defect function nonzero extent.

The broadening effect from spherical defects and surface irregularities has been extensively covered in the literature; for example, a thorough discussion of these types of broadening has been given by Hernandez (1966, 1986) and by Hays and Roble (1971). A treatment of misalignment broadening due to a lack of parallelism of the etalon plates has been done by Gupta and Prasad (1991). Atherton et al. (1981)

cover characterization of these defects including their corresponding defect finesses. These plate defect characteristics are summarized in Table A.2. The expression for standard deviation of the Gaussian defect distribution can be easily arrived upon by combining the definition of standard deviation for a Gaussian ($\sigma \equiv \frac{\delta_H}{\sqrt{2 \ln 2}}$, where $\delta_H \equiv$ half-width at half-maximum (HWHM)) with the HWHM expression for a defect ($\delta_{H_D} = \frac{\nu_{sr}}{2N_D}$, where the defect finesse $N_D = N_{D_G}$ for this case) given by Hays and Roble (1971). The formulation for φ_{max} for a spherical defect can be obtained by noting that $\varphi_{max} = \delta_{H_D}$ for a rectangular distribution. And, finally, the φ_{max} relationship given for the parallelism error is obvious from the development presented in Gupta and Prasad (1991).

A particular defect distribution function may apply to several very different types of surface defects. For example, Sloggett (1984) discusses how a rectangular distribution may also be used to characterize uniformly distributed random defects, a parabolic defect, and even a parallelism error in plates used with a square aperture. In addition, other types of plate defects may exist which are best described using distribution functions other than those given above. For example, a surface defect of sinusoidal form might result from certain polishing techniques (Vaughan, 1989); its corresponding defect distribution function would be an inverted semiellipse having peaks at the extremes of phase difference.

Equations A.16 and A.17 show how the defect distribution function $D(\varphi)$ can be used to arrive upon the etalon transmission function $E(\Psi)$. In particular, the integral in equation A.16 may be solved analytically for several forms of the defect function (Hernandez, 1966, 1986; Hays and Roble, 1971), or it may simply be evaluated numerically. Alternatively, the Airy function can be numerically convolved with the defect distribution function (equation A.17) using the discrete Fast Fourier Trans-

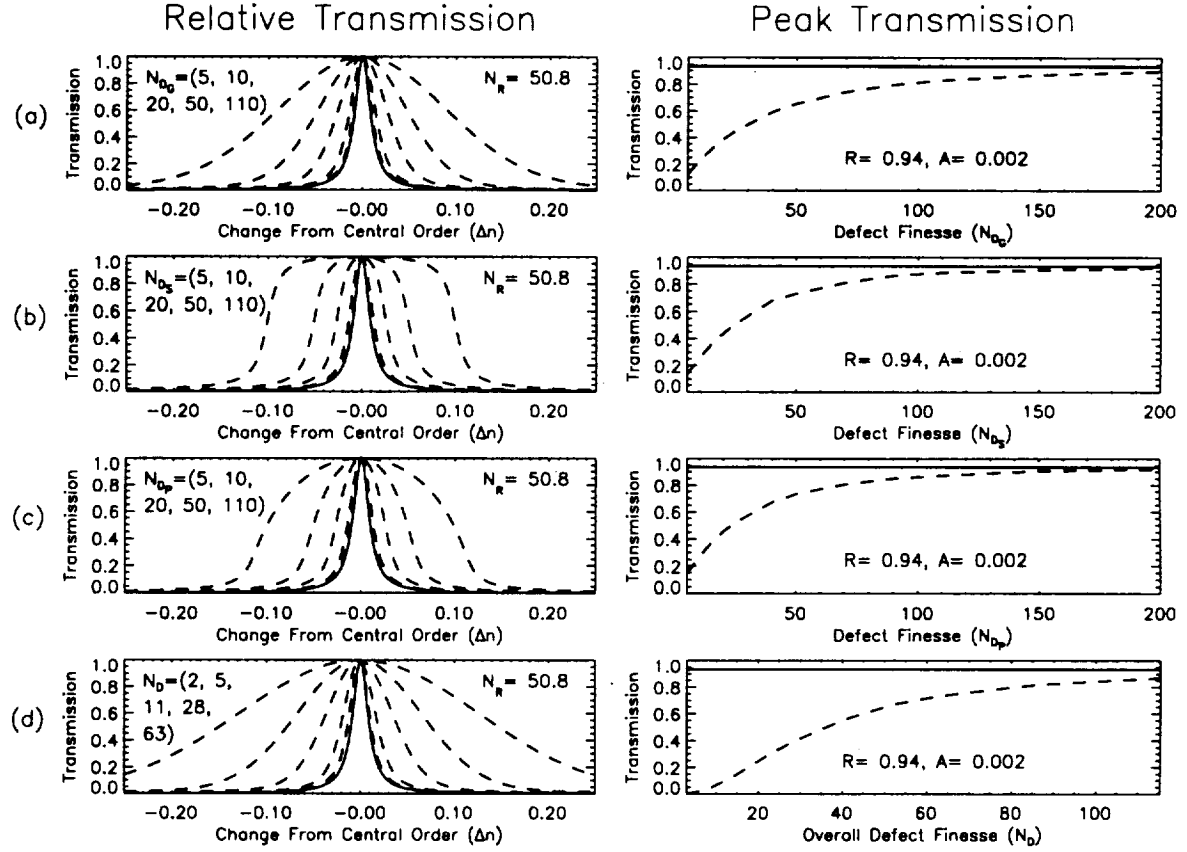


Figure A.7: Etalon transmission characteristics in the presence of plate defects. Plots on the left side of the Figure show passband transmission variation near the central order normalized to the peak transmission for several values of plate defect finesse; whereas the plots on the right depict the peak transmission dependence on plate defect finesse. The solid curves represent plates without defects. (a) represents a Gaussian defect distribution function for a random surface imperfection. (b) corresponds to a rectangular defect distribution function for a spherical plate defect. A semielliptical defect function was used in (c) to represent a plate parallelism error. And (d) shows the combined effect of the defects represented in (a) through (c) occurring simultaneously.

form (FFT) (Churchill, 1972). This later approach has been implemented in this research upon its validation through comparisons among the different techniques. The Convolution Theorem (Churchill, 1972; Press et al., 1990) may be interpreted as saying that the Fourier transform of a convolution is equal to the product of the individual Fourier transforms. Accordingly, equation A.17 may be expressed as

$$E(\Psi) = \mathcal{A}(\Psi) * D(\Psi) = FFT^{-1}\{FFT\{\mathcal{A}(\Psi)\}FFT\{D(\Psi)\}\}. \quad (\text{A.19})$$

For the realistic case of several plate imperfections existing simultaneously, the defect distribution function can be formulated by convolving all of the individual defect functions together. This can be expressed as

$$D(\Psi) = D_G(\Psi) * D_S(\Psi) * D_P(\Psi) \quad (\text{A.20})$$

for the three main defects discussed earlier. Figure A.7 illustrates the effects of plate imperfections on the characteristics of the etalon transmission function. The reflectivity and absorption/scattering of the plates are assumed to be .94 ($N_R \approx 50$) and .002, respectively. The solid curves represent plates without defects; for example, the transmission function with reflectivity broadening only (Airy function, equation A.1), and peak transmission reduced only from absorption/scattering of the plates (equation A.2). Figures A.7(a)-(c) show how Gaussian, rectangular, and semielliptical defect distribution functions, respectively, affect the etalon transmission function when acting alone for different values of defect finesse. Whereas, Figure A.7(d) depicts the combined effect of the first three defects occurring simultaneously, with equal individual defect finesse of same value as the corresponding curves in (a) through (c) of the Figure (and thus yielding an overall defect finesse given by equation A.18). As the plate quality gets better ($N_D \rightarrow N_D \gg N_R$), the etalon transfer function approaches the Airy function; conversely, as the plate quality gets

poorer ($N_D \rightarrow N_D \ll N_R$), the etalon transfer function gets broadened into a shape resembling the defect function itself.

In addition to plate imperfections, an aperture in the optical system will also produce broadening of the FPI transfer function. The radiation passing through an etalon will cover some finite solid angle, which is usually defined by an aperture placed in the fringe plane. The most common configuration involves a fixed, circular aperture centered on the optical axis. And, as with plate defects, the effect will be to broaden the FPI instrumental transmission function. This can be represented mathematically by

$$T(\Psi) = E(\Psi) * L(\Psi), \quad (\text{A.21})$$

where $L(\Psi)$ is the aperture function, $E(\Psi)$ is the etalon transmission function as given by equation A.17, and $T(\Psi)$ corresponds to the FPI instrumental transmission function; it expresses the transmission of the interferometer with imperfect etalon plates and a finite sized aperture. As with the function used to represent a spherical defect of etalon plates, the aperture function can also be expressed as a rectangular function (Hernandez, 1986; Hays and Roble, 1971). However, the half-width at half-maximum (HWHM) of the aperture function is related to the optical geometry of the FPI; and for a circular aperture in the fringe plane it can be given by

$$\delta_{H_L} = \frac{\nu d_L^2}{(4f)^2}, \quad (\text{A.22})$$

where ν is the wavenumber of incident radiation, d_L is the diameter of the circular aperture, and f is the focal length of the fringe forming lens. As mentioned by Hays and Roble (1971), a true finesse cannot be assigned to aperture function broadening since (unlike the Airy and defect functions) its HWHM δ_{H_L} is not a function of the

plate separation t . However, it is still useful to define a pseudo aperture finesse by

$$N_L = \frac{\nu_{f_{sr}}}{2\delta_{H_L}}; \quad (\text{A.23})$$

which for the circular aperture system can be given by $N_L = \frac{8f^2\nu_{f_{sr}}}{\nu d_L^2} \approx \frac{4f^2}{\nu t d_L^2}$. The effective resolution (or full-width at half-maximum (FWHM)) of the FPI is a function of the reflectivity, defect, and aperture finesse; this can be given approximately by

$$\delta_{F_E} = \frac{\nu_{f_{sr}}}{N_E}, \quad (\text{A.24})$$

where N_E is an “effective finesse” that can be defined by

$$\frac{1}{N_E^2} = \frac{1}{N_R^2} + \frac{1}{N_D^2} + \frac{1}{N_L^2}. \quad (\text{A.25})$$

The luminosity-resolution product (LRP) (Jacquinot, 1954) is usually optimum when $N_R \approx N_D \approx N_L$ (Atherton et al., 1981). This is usually accomplished by first obtaining plates with a minimum of imperfections (maximizing N_D), then applying a reflectivity coating which equates N_R to N_D , and then implementing an aperture/detector system such that $N_L \approx N_D$. However, depending upon the application, it may be desirable to alter this finesse selection procedure and allow N_R , N_D , and N_L to have different values. For example, if the expected signal level is of concern and if the corresponding resolution is higher than necessary, then N_R and N_L can be made less than N_D . Conversely, to increase resolution at the expense of lowering the signal, one can specify N_R and N_L to be greater than N_D . However, it should be noted that as $N_R \approx N_D \rightarrow N_R \gg N_D$ there would be significantly more reduction in throughput than would be gained in overall instrument finesse.

APPENDIX B

ERROR ANALYSIS: SENSITIVITIES TO SYSTEMATIC PERTURBATIONS

This appendix contains supplemental material for the error analysis discussed within Section 5.6. The following tables summarize the potential impact on retrieved integrated ozone amounts from systematic offsets of certain forward model parameters for reference atmospheres used in this study.

Results of Error Analysis: 2) mid-latitude summer (45N July)		
Effect on retrieved integrated O_3 amounts		
Perturbation	Accuracy (%)	
	(systematic errors)	
	Troposphere	Total
H_2O continuum model		
without H_2O continuum	24.43	3.856
Cs & Cn +10%	-2.318	-0.3659
surface emissivity		
ϵ_s -1%	-9.704	-1.495
surface temperature		
T_s +1K	17.30	2.666
temperature profile		
T+1K all levels	4.838	-3.607
interfering species		
CO_2 +1% all levels	-0.02581	-0.01027
CH_4 +10% all levels	0.0001292	-0.0001588
H_2O +10% all levels	-4.049	-0.7317
N_2O +30% all levels	-0.0005321	6.873e-05
SO_2 +30% all levels	-0.0001197	1.531e-05

Table B.1: The effect of certain systematic perturbations on retrieved integrated ozone amounts for a mid-latitude summer atmosphere.

Results of Error Analysis: 3) mid-latitude winter (45N Jan)		
	Effect on retrieved integrated O_3 amounts	
Perturbation	Accuracy (%) (systematic errors)	
	Troposphere	Total
H_2O continuum model		
without H_2O continuum	3.666	0.1076
Cs & Cn +10%	-0.3633	-0.01062
surface emissivity		
ϵ_s -1%	-9.641	-0.2660
surface temperature		
T_s +1K	20.08	0.5540
temperature profile		
T+1K all levels	-3.995	-0.8588
interfering species		
CO_2 +1% all levels	-0.01034	-0.003279
CH_4 +10% all levels	0.0002633	-6.712e-05
H_2O +10% all levels	-0.5218	-0.02885
N_2O +30% all levels	-0.0002198	5.256e-05
SO_2 +30% all levels	2.605e-05	2.116e-06

Table B.2: The effect of certain systematic perturbations on retrieved integrated ozone amounts for a mid-latitude winter atmosphere.

Results of Error Analysis: 4) subarctic summer (60N July)		
	Effect on retrieved integrated O_3 amounts	
Perturbation	Accuracy (%) (systematic errors)	
	Troposphere	Total
H_2O continuum model		
without H_2O continuum	22.04	1.908
Cs & Cn +10%	-2.126	-0.1841
surface emissivity ϵ_s -1%	-11.28	-0.9499
surface temperature T_s +1K	21.10	1.777
temperature profile T+1K all levels	0.8645	-3.828
interfering species		
CO_2 +1% all levels	-0.02320	-0.007093
CH_4 +10% all levels	0.0004479	-0.0001548
H_2O +10% all levels	-3.604	-0.3845
N_2O +30% all levels	-0.0001661	4.002e-05
SO_2 +30% all levels	-1.263e-05	1.959e-05

Table B.3: The effect of certain systematic perturbations on retrieved integrated ozone amounts for a subarctic summer atmosphere.

Results of Error Analysis: 5) subarctic winter (60N Jan)		
	Effect on retrieved integrated O_3 amounts	
Perturbation	Accuracy (%)	
	(systematic errors)	
	Troposphere	Total
H_2O continuum model		
without H_2O continuum	0.4382	0.0009927
Cs & Cn +10%	-0.04363	-0.0001334
surface emissivity		
ϵ_s -1%	-6.913	0.02194
surface temperature		
T_s +1K	16.14	-0.05098
temperature profile		
T+1K all levels	-5.797	3.596
interfering species		
CO_2 +1% all levels	-0.001987	-0.001434
CH_4 +10% all levels	0.0001559	-5.254e-05
H_2O +10% all levels	-0.04769	-0.002093
N_2O +30% all levels	2.226e-05	-2.203e-05
SO_2 +30% all levels	1.546e-06	-7.845e-06

Table B.4: The effect of certain systematic perturbations on retrieved integrated ozone amounts for a subarctic winter atmosphere.

Results of Error Analysis: 6) U.S. standard (1976)		
	Effect on retrieved integrated O_3 amounts	
Perturbation	Accuracy (%) (systematic errors)	
	Troposphere	Total
H_2O continuum model		
without H_2O continuum	15.31	1.470
Cs & Cn +10%	-1.503	-0.1441
surface emissivity		
ϵ_s -1%	-13.16	-1.241
surface temperature		
T_s +1K	24.44	2.305
temperature profile		
T+1K all levels	-2.057	-4.307
interfering species		
CO_2 +1% all levels	-0.02481	-0.008861
CH_4 +10% all levels	0.0003253	-0.0001757
H_2O +10% all levels	-2.450	-0.2948
N_2O +30% all levels	0.0001823	2.717e-05
SO_2 +30% all levels	0.0001796	-6.827e-06

Table B.5: The effect of certain systematic perturbations on retrieved integrated ozone amounts for the U.S. standard (1976) atmosphere.

APPENDIX C

ERROR ANALYSIS: OVERALL ERROR BUDGETS

This appendix contains supplemental material for the error analysis discussed within Section 5.6. The following tables summarize estimates of overall errors in retrieved integrated ozone amounts based upon the assumed systematic and random uncertainties for reference atmospheres used in this study.

Results of Error Analysis: 2) mid-latitude summer (45N July)				
Effect on retrieved integrated O_3 amounts				
Perturbation	Accuracy (%) (systematic errors)		Precision (%) (random errors)	
	Troposphere	Total	Troposphere	Total
H_2O continuum model				
Cs +10%	-1.675	-0.2638		
Cn +10%	-0.6466	-0.1027		
surface emissivity $\epsilon_s \pm 1\%$			2.215	0.4450
signal calibration all channels +1%	11.95	0.2213		
surface temperature $T_s \pm 1K$			3.948	0.7932
temperature profile $T \pm 2k$ one level at a time			1.238	0.9294
interfering species				
$H_2O \pm 30\%$ one level at a time			0.9955	0.2027
$CO_2 \pm 5\%$ one level at a time			0.01317	0.003535
spectral line parameters				
H_2O halfwidths +5%	-0.01318	-0.02858		
O_3 halfwidths +5%	-8.885	-3.000		
H_2O line strengths +10%	0.03848	-0.08861		
O_3 line strengths +1%	-1.349	-0.8957		
Error impact (rss)				
forward model uncertainties	15.12	3.278	4.798	1.316
measurement noise			2.328	1.218
null-space error			5.582	2.781
total (rss) errors	15.12	3.278	7.730	3.311

Table C.1: Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a mid-latitude summer atmosphere.

Results of Error Analysis: 3) mid-latitude winter (45N Jan)				
Effect on retrieved integrated O_3 amounts				
Perturbation	Accuracy (%) (systematic errors)		Precision (%) (random errors)	
	Troposphere	Total	Troposphere	Total
H_2O continuum model				
Cs +10%	-0.1671	-0.004771		
Cn +10%	-0.1966	-0.005808		
surface emissivity $\epsilon_s \pm 1\%$			2.213	0.5462
signal calibration all channels +1%	8.445	-0.1085		
surface temperature $T_s \pm 1K$			4.609	1.137
temperature profile $T \pm 2k$ one level at a time			0.7924	0.5695
interfering species				
$H_2O \pm 30\%$ one level at a time			0.1337	0.03042
$CO_2 \pm 5\%$ one level at a time			0.004020	0.0009045
spectral line parameters				
H_2O halfwidths +5%	0.0006142	-0.002767		
O_3 halfwidths +5%	-8.273	-1.716		
H_2O line strengths +10%	0.02382	-0.01300		
O_3 line strengths +1%	-1.189	-0.7040		
Error impact (rss)				
forward model uncertainties	11.94	1.987	5.175	1.385
measurement noise			2.517	1.284
null-space error			6.426	3.154
total (rss) errors	11.94	1.987	8.627	3.681

Table C.2: Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a mid-latitude winter atmosphere.

Results of Error Analysis: 4) subarctic summer (60N July)				
Effect on retrieved integrated O_3 amounts				
Perturbation	Accuracy (%) (systematic errors)		Precision (%) (random errors)	
	Troposphere	Total	Troposphere	Total
H_2O continuum model				
Cs +10%	-1.436	-0.1239		
Cn +10%	-0.6929	-0.06028		
surface emissivity $\epsilon_s \pm 1\%$			2.363	0.4212
signal calibration all channels +1%	11.98	-0.5733		
surface temperature $T_s \pm 1K$			4.420	0.7877
temperature profile $T \pm 2k$ one level at a time			0.9182	0.8352
interfering species				
$H_2O \pm 30\%$ one level at a time			0.8793	0.1518
$CO_2 \pm 5\%$ one level at a time			0.008623	0.002129
spectral line parameters				
H_2O halfwidths +5%	-4.776e-05	-0.01980		
O_3 halfwidths +5%	-9.227	-2.630		
H_2O line strengths +10%	0.07047	-0.06774		
O_3 line strengths +1%	-1.260	-0.8150		
Error impact (rss)				
forward model uncertainties	15.31	2.932	5.171	1.232
measurement noise			2.360	1.225
null-space error			5.939	2.822
total (rss) errors	15.31	2.932	8.224	3.317

Table C.3: Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a subarctic summer atmosphere.

Results of Error Analysis: 5) subarctic winter (60N Jan)				
	Effect on retrieved integrated O_3 amounts			
Perturbation	Accuracy (%) (systematic errors)		Precision (%) (random errors)	
	Troposphere	Total	Troposphere	Total
H_2O continuum model				
Cs +10%	-0.008697	-3.080e-05		
Cn +10%	-0.03506	-6.291e-05		
surface emissivity $\epsilon_s \pm 1\%$			2.083	0.5924
signal calibration all channels +1%	5.007	0.9888		
surface temperature $T_s \pm 1K$			4.863	1.383
temperature profile $T \pm 2k$ one level at a time			1.026	0.8166
interfering species				
$H_2O \pm 30\%$ one level at a time			0.02406	0.006576
$CO_2 \pm 5\%$ one level at a time			0.001325	0.0004652
spectral line parameters				
H_2O halfwidths +5%	0.0001325	-0.0003161		
O_3 halfwidths +5%	-5.522	-1.763		
H_2O line strengths +10%	0.004202	-0.001945		
O_3 line strengths +1%	-0.8424	-0.6742		
Error impact (rss)				
forward model uncertainties	7.549	2.235	5.389	1.712
measurement noise			2.893	1.290
null-space error			8.313	3.303
total (rss) errors	7.549	2.235	10.32	3.945

Table C.4: Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to a subarctic winter atmosphere.

Results of Error Analysis: 6) U.S. standard (1976)				
Effect on retrieved integrated O_3 amounts				
Perturbation	Accuracy (%) (systematic errors)		Precision (%) (random errors)	
	Troposphere	Total	Troposphere	Total
H_2O continuum model				
Cs +10%	-0.8781	-0.08415		
Cn +10%	-0.6261	-0.06031		
surface emissivity $\epsilon_s \pm 1\%$			2.894	0.4772
signal calibration all channels +1%	12.76	-0.2557		
surface temperature $T_s \pm 1K$			5.376	0.8864
temperature profile $T \pm 2k$ one level at a time			1.116	0.9058
interfering species				
$H_2O \pm 30\%$ one level at a time			0.5966	0.09579
$CO_2 \pm 5\%$ one level at a time			0.01057	0.002926
spectral line parameters				
H_2O halfwidths +5%	-0.01611	-0.01684		
O_3 halfwidths +5%	-10.03	-2.503		
H_2O line strengths +10%	0.02241	-0.05835		
O_3 line strengths +1%	-1.358	-0.8295		
Error impact (rss)				
forward model uncertainties	16.38	2.779	6.236	1.358
measurement noise			2.793	1.272
null-space error			6.808	2.868
total (rss) errors	16.38	2.779	9.656	3.421

Table C.5: Overall errors in retrieved integrated ozone amounts corresponding to application of the proposed instrumentation to the U.S. standard (1976) atmosphere.

BIBLIOGRAPHY

BIBLIOGRAPHY

- [Abr.1] Abreu, V.J., and Skinner, W.R., "Inversion of Fabry-Perot CCD images: use in Doppler shift measurements," *Appl. Opt.*, vol.28, 1989, pp.3382-3386.
- [And.1] Anderson, G.P., Clough, S.A., Kneizys, F.X., Chetwynd, J.H., and Shettle, E.P., **AFGL Atmospheric Constituent Profiles (0-120 km)** AFGL-TR-86-0110, 1986.
- [Arm.1] Armstrong, B.H., "Spectrum Line Profiles: The Voigt Function," *J. Quant. Spectrosc. Radiat. Transfer*, vol.7, 1967, pp.61-88.
- [Ath.1] Atherton, P.D., Reay, N.K., and Ring, J., "Tunable Fabry-Perot filters," *Opt. Eng.*, vol.20, 1981, pp.806-814.
- [Bak.1] Baker, R.L., Mauldin, L.E., III, and Russell, J.M., III, "Design and performance of the halogen occultation experiment (HALOE) remote sensor," *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol.685, 1986, pp.181-191.
- [Bar.1] Barmore, F.E., "High Resolution Observations of the 6300 A Oxygen line in the Dayglow," *Planet. Space Sci.*, vol.25, 1977, pp.185-191.
- [Bar.1] Bartia, P.K., "Precisions and Accuracies for TOMS," NASA Goddard Space Flight Center, private communication, September, 1993.
- [Ben.1] Bens, A.R., Cogger, L.L., and Shepherd, G.G., "Upper Atmospheric Temperatures From Doppler Line Widths-III," *Planet. Space Sci.*, vol.13, 1965, pp.551-563.
- [Bla.1] Blamont, J.E., and Luton, J.M., "Geomagnetic Effect on the Neutral Temperature of the F Region during the Magnetic Storm of September 1969," *J. Geophys. Res.*, vol.77, 1972, pp.3534-3556.
- [Bor.1] Born, M., and Wolf, E., **Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light**. Oxford: Pergamon, 5th ed. 1975.
- [Bra.1] Bradley, L.C., and Kuhn, H., "Spectrum of helium-3," *Nature*, vol.162, 1948, pp.412-413.

- [Bur.1] Burch, D.E., and Alt, R.L., **Continuum Absorption by H_2O in the 700-1200 cm^{-1} and 2400-2800 cm^{-1} Windows.** AFGL-TR-84-0128, 1984.
- [Bur.2] Burch, D.E., "Continuum Absorption by Atmospheric H_2O ," Proc. Soc. Photo-Instrum. Eng., vol.277, 1981, pp.28-39.
- [Bur.1] Burleigh Instruments, Inc., **Fabry-Perot Interferometers.** Burleigh technical memo, NewYork.
- [Can.1] Cannell, D.S., and Benedek, G.B., "Brillouin spectrum of xenon near its critical point," Phys. Rev. Lett., vol.25, 1970, pp.1157-1161.
- [Cha.1] Chantry, G.W., "The use of Fabry-Perot Interferometers, Etalons and Resonators at Infrared and Longer Wavelengths-an Overview," J. Phys. E: Sci. Instrum., vol.15, 1982, pp.3-8.
- [Chu.1] Churchill, R.V., **Operational Mathematics.** New York: McGraw-Hill, 1972.
- [Clo.1] Clough, S.A., Kneizys, F.X., Davies, R., Gamache, R., and Tipping, R., "Theoretical Line Shape for H_2O Vapor; Application to the Continuum," in **Atmospheric Water Vapor.** eds. A. Deepak, T. D. Wilkerson, and L.H. Ruhnke, NewYork: Academic, 1980, pp.25-46.
- [Coc.1] Cocks, T.D., Creighton, D.F., and Jacka, F., "Application of a dual Fabry-Perot spectrometer for daytime airglow studies," J. Atmos. Terr. Phys., vol.42, 1980, pp.499-511.
- [Dae.1] Daehler, M., "A Twelve-Channel Multiple-Interferometer Fabry-Perot Spectrometer," Appl. Opt., vol.9, 1970, pp.2529-2534.
- [Dae.2] Daehler, M., and Roesler, F.L., "High Contrast in a Polyetalon Fabry-Perot Spectrometer," Appl. Opt., vol.7, 1968, pp.1240-1241.
- [Del.1] Del Piano, V.N., and Quesada, A.F., "Transmission Characteristics of Fabry-Perot Interferometers and a Related Electrooptic Modulator," Appl. Opt., vol.4, 1965, pp.1386-1390.
- [Dra.1] Drayson, S.R., "Rapid Computation of the Voigt Profile," J. Quant. Spectrosc. Radiat. Transfer, vol.16, 1976, pp.611-614.
- [Dra.2] Drayson, S.R., "Atmospheric transmission in the CO_2 bands between 12μ and 18μ ," Appl. Opt., vol.5, 1966, pp.385-391.
- [Far.1] Farman, J.C., Gardiner, B.G., and Shanklin, J.D., "Large losses of total ozone in Antarctica reveal seasonal ClO_x/NO_x interaction," Nature, vol.315, 1985, pp.207-210.

- [Fis.1] Fishman, J., "Probing Planetary Pollution From Space," *Environ. Sci. Technol.*, vol.25, 1991, pp.612-621.
- [Fis.2] Fishman, J., Watson, C.E., Larsen, J.C., and Logan, J.A., "Distribution of Tropospheric Ozone Determined from Satellite Data," *J. Geophys. Res.*, vol.95, 1990, pp.3599-3617.
- [Fis.3] Fishman, J., *Tropospheric Ozone from Satellite Total Ozone Measurements in Tropospheric Ozone: regional and global scale interactions*. ed. I.S.A. Isaksen, Boston: Dordrecht, 1988, pp.111-123
- [Fis.4] Fishman, J., and Browell, E.V., "Comparison of satellite total ozone measurements with the distribution of tropospheric ozone obtained by an airborne UV-DIAL system over the Amazon Basin," *Tellus*, vol.40B, 1988, pp.393-407.
- [Fis.5] Fishman, J., Vukovich, F.M., Cahoon, D.R., and Shipham, M.C., "The Characterization of an Air Pollution Episode Using Satellite Total Ozone Measurements," *J. Clim. Appl. Met.*, vol.26, 1987, pp.1638-1654.
- [Fis.6] Fishman, J., and Larsen, J., "Distribution of Total Ozone and Stratospheric Ozone in the Tropics: Implications for the Distribution of Tropospheric Ozone," *J. Geophys. Res.*, vol.92, 1987, pp.6627-6634.
- [Fis.7] Fishman, J., Minnis, P., and Reichle, H.G., Jr., "Use of Satellite Data to Study Tropospheric Ozone in the Tropics," *J. Geophys. Res.*, vol.91, 1986, pp.14,451-14,465.
- [Fla.1] Flaud, J.-M., Camy-Peyret, C., Rinsland, C.P., Devi, V.M., Smith, M.A.H., and Goldman A., "Improved line parameters for ozone bands in the 10- μ m spectral region," *Appl. Opt.*, vol.29, 1990, pp.3667-3671.
- [Fre.1] Frederick, J.E., Cebula, R.P., and Heath, D.F., "Instrument characterization for the detection of long-term changes in stratospheric ozone: an analysis of the SBUV/2 radiometer," *J. Atmos. Oceanic Technol.*, vol.3, 1986, pp.472-480.
- [Gal.1] Gallery, W.O., Kneizys, F.X., and Clough, S.A., **Air Mass Computer Program for Atmospheric Transmittance/Radiance: FSCATM** AFGL-TR-83-0065, 1983.
- [Gil.1] Gille, J.C., and Russell, J.M., III, "The Limb Infrared Monitor of the Stratosphere: Experiment Description, Performance, and Results," *J. Geophys. Res.*, vol.89, 1984, pp.5125-5140.
- [Gra.1] Grant, W.B., "Water vapor absorption coefficients in the 8-13- μ m spectral region: a critical review," *Appl. Opt.*, vol.29, 1990, pp.451-462.

- [Gra.1] Grant, W.B., ed., **Ozone Measuring Instruments for the Stratosphere**. Vol. 1 of Collected Works in Optics, Optical Society of America, Washington D.C., 1989.
- [Gre.1] Gregg, W.W., Ardanuy, P.E., Braun, W.C., and Vallette, B.J., **Analysis of Error in TOMS Total Ozone as a Function of Orbit and Attitude Parameters**. NASA Contractor Report 4361, 1991.
- [Gup.1] Gupta, R., and Prasad, C.D., "Instrumental broadening caused by the misalignment function in a Fabry-Perot etalon assembly," *Appl. Opt.*, vol.30, 1991, pp.373-375.
- [Han.1] Hanel, R.A., Schlachman, B., Rogers, D., and Vanous, D., "Nimbus 4 Michelson interferometer," *Appl. Opt.*, vol.10, 1971, pp.1376-1382.
- [Hay.1] Hays, P.B., "Space-flight Preparation Impact on Etalon Defect Finesse," The University of Michigan, UARS science team meeting in Virginia Beach, VA, private communication, March, 1993.
- [Hay.2] Hays, P.B., Wang, J., "Image plane detector for Fabry-Perot interferometers: physical model and improvement with anticoincidence detection," *Appl. Opt.*, vol.30, 1991, pp.3100-3107.
- [Hay.3] Hays, P.B., "Circle to Line Interferometer Optical System," *Appl. Opt.*, vol.29, 1990, pp.1482-1489.
- [Hay.4] Hays, P.B., "High-Resolution Optical Measurements of Atmospheric Wind from Space. 1: Lower Atmosphere Molecular Absorption," *Appl. Opt.*, vol.21, 1982, pp.1136-1141.
- [Hay.5] Hays, P.B., Killeen, T.L., and Kennedy, B.C., "The Fabry-Perot Interferometer on Dynamics Explorer," *Space Sci. Instrum.*, vol.5, 1981, pp.395-416.
- [Hay.6] Hays, P.B., Roble, R.G., "A Technique for Recovering Doppler Line Profiles from Fabry-Perot Interferometer Fringes of Very Low Intensity," *Appl. Opt.*, vol.10, 1971, pp.193-200.
- [Hea.1] Heath, D.F., Krueger, A.J., Roeder, H.A., and Henderson, B.D., "The solar backscatter ultraviolet and total ozone mapping spectrometer (SBUV/TOMS) for Nimbus G," *Opt. Eng.*, vol.14, 1975, pp.323-331.
- [Her.1] Hernandez, G., **Fabry-Perot interferometers**. NewYork: Cambridge, 1986.
- [Her.2] Hernandez, G., "Analytical description of a Fabry-Perot Photoelectric Spectrometer," *Appl. Opt.*, vol.5, 1966, pp.1745-1748.

- [Hir.1] Hirschberg, J.G., and Platz, P., "A Multichannel Fabry-Perot Interferometer," *Appl. Opt.*, vol.4, 1965, pp.1375-1381.
- [Hou.1] Houghton, J.T., Taylor, F.W., and Rodgers, C.D., **Remote sounding of atmospheres**. New York: Cambridge, 1984.
- [Hou.2] Houghton, J.T., "The Meteorological Significance of Remote Measurements of Infra-red Emission from Atmospheric Carbon Dioxide," *Quart. J. Roy. Met. Soc.*, vol.87, 1961, pp.102-104.
- [Hou.1] Houston, W.V., "A compound interferometer for fine structure work," *Phys. Rev.*, vol.29, 1965, pp.478-484.
- [Jac.1] Jacquinet, P., "The Luminosity of Spectrometers with Prisms, Gratings, or Fabry-Perot Etalons," *J. Opt. Soc. Am.*, vol.44, 1954, pp.761-765.
- [Khr.1] Khrigian, A.Kh., **The Physics of Atmospheric Ozone**. Jerusalem: Keterpress, 1975.
- [Kil.1] Killeen, T.L., and Hays, P.B., "Doppler line profile analysis for a multichannel Fabry-Perot interferometer," *Appl. Opt.*, vol.23, 1984, pp.612-620.
- [Kil.2] Killeen, T.L., Kennedy, B.C., Hays, P.B., Symanow, D.A., and Ceckowski, D.H., "Image plane detector for the Dynamics Explorer Fabry-Perot interferometer," *Appl. Opt.*, vol.22, 1983, pp.3503-3513.
- [Kuh.1] Kuhn, W.R., and London, J., "Infrared radiative cooling in the middle atmosphere (30-110 km)," *J. Atmos. Sci.*, vol.26, 1969, pp.189-204.
- [Log.1] Logan, J.A., Prather, M.J., Wofsy, S.C., and McElroy, M.B., "Tropospheric Chemistry: A Global Perspective," *J. Geophys. Res.*, vol.86, 1981, pp.7210-7254.
- [Ma.1] Ma, X.-L., Smith, W.L., and Woolf, H.M., "Total Ozone from NOAA Satellites—A Physical Model for Obtaining Measurements with High Spatial Resolution," *J. Clim. Appl. Met.*, vol.23, 1984, pp.1309-1314.
- [Mac.1] Mack, J.E., McNutt, D.P., Roesler, F.L., and Chabbal, R., "The PEP-SIOS Purely Interferometric High-Resolution Scanning Spectrometer. I. The Pilot Model," *Appl. Opt.*, vol.2, 1963, pp.873-884.
- [McC.1] McCormick, M.P., "SAGE II: an overview," *Adv. Space Res.*, vol.7, 1987, pp.219-226.
- [McC.2] McCormick, M.P., Hamill, P., Pepin, T.J., Chu, W.P., Swissler, T.J., and McMaster, L.R., "Satellite studies of the stratospheric aerosol," *Bull. Am. Meteorol. Soc.*, vol.60, 1979, pp.1038-1046.

- [Pen.1] Penkett, S.A., "Increased tropospheric ozone," *Nature*, vol.332, 1988, pp.204-205.
- [Pla.1] Planet, W.G., Crosby, D.S., Lienesch, J.H., and Hill, M.L. , "Determination of Total Ozone Amount from TIROS Radiance Measurements," *J. Clim. Appl. Met.*, vol.23, 1984, pp.308-316.
- [Pre.1] Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterling, W.T., **Numerical Recipes The Art of Scientific Computing (FORTRAN version)**. NewYork: Cambridge, 1990.
- [Ree.1] Rees, D., Rounce, P.A., McWhirter, I., Scott, A.F.D., Greenaway, A.H., and Towlson, W., "Observations of atmospheric absorption lines from a stabilised balloon platform and measurements of stratospheric winds," *J. Phys. E: Sci. Instrum.*, vol.15, 1982, pp.191-206.
- [Ree.2] Rees, D., McWhirter, I., Hays, P.B., and Dines, T., "A Stable, Rugged, Capacitance-Stabilized Piezoelectric Scanned Fabry-Perot Etalon," *J. Phys. E*, vol.14, 1981, pp.1320-1328.
- [Roc.1] Roche, A.E., and Kumer, J.B., "Cryogenic limb array etalon spectrometer (CLAES): experiment overview," *Proc. Soc. Photo-Opt. Instrum. Eng.*, vol.973, 1989, pp.324-334.
- [Rod.1] Rodgers, C.D., "Characterization and Error Analysis of Profiles Retrieved From Remote Sounding Measurements," *J. Geophys. Res.*, vol.95, 1990, pp.5587-5595.
- [Rod.2] Rodgers, C.D., "A General Error Analysis for Profile Retrieval," in **RSRM '87: Advances in Remote Sensing Retrieval Methods**. A. Deepak, H.E. Fleming, and J.S. Theon (Eds.), Virginia: A. Deepak Pub., 1989.
- [Rod.3] Rodgers, C.D., "Retrieval Theory," in **Remote sounding of atmospheres**. Houghton, J.T., Taylor, F.W., and Rodgers, C.D., NewYork: Cambridge, 1984.
- [Rod.4] Rodgers, C.D., "Retrieval of Atmospheric Temperature and Composition from Remote Measurements of Thermal Radiation," *Rev. Geophys. and Space Phys.*, 14, 1976, pp.609-624.
- [Roe.1] Roesler, F., "Fabry-Perot instruments for astronomy," in **Methods of Experimental Physics**. vol.12, part A ed. N. Carleton, NewYork: Academic, 1974, pp.531-568
- [Roe.2] Roesler, F.L., "Effects of Plate Defects in a Polyetalon Fabry-Perot Spectrometer," *Appl. Opt.*, vol.8, 1969, pp.829-831.

- [Rot.1] Rothman, L.S., "The HITRAN Molecular Database: Enhancements for Remote Sensing," in **Optical Remote Sensing of the Atmosphere**. vol.5 of 1993 Technical Digest Series, Optical Society of America, Washington D.C., 1993.
- [Rot.2] Rothman, L.S., Gamache, R.R., Goldman A., Brown, L.R., Toth, R.A., Pickett, H.M., Poynter, R.L., Flaud, J.-M., Camy-Peyret, C., Barbe, A., Husson, N., Rinsland, C.P., and Smith, M.A.H., "The HITRAN database: 1986 edition," *Appl. Opt.*, vol.26, 1987, pp.4058-4097.
- [Rus.1] Russell, J.M., III, Gordley, L.L., Park, J.H., Drayson, S.R., Hesketh, W.D., Cicerone, R.J., Tuck, A.F., Frederick, J.E., Harries, J.E., and Crutzen, P.J., "The Halogen Occultation Experiment," *J. Geophys. Res.*, vol.98, 1993, pp.10,777-10,797.
- [Rus.2] Russell, J.M., III, and Drayson, S.R., "The Inference of Atmospheric Ozone Using Satellite Horizon Measurements in the 1042 cm^{-1} Band," *J. Atmos. Sci.*, vol.29, 1972, pp.376-390.
- [Sei.1] Seinfeld, J.H., **Atmospheric Chemistry and Physics of Air Pollution**. New York: Wiley, 1986.
- [Ski.1] Skinner, W.R., Hays, P.B., and Abreu, V.J., "Optimization of a Triple Etalon Interferometer," *Appl. Opt.*, vol.26, 1987, pp.2817-2827.
- [Slo.1] Sloggett, G.J., "Fringe Broadening in Fabry-Perot Interferometers," *Appl. Opt.*, vol.23, 1984, pp.2427-2432.
- [Sto.1] Stolarski, R.S., Krueger, A.J., Schoeberl, M.R., McPeters, R.D., Newman, P.A., and Alpert, J.C., "Nimbus 7 satellite measurements of the springtime Antarctic ozone decrease," *Nature*, vol.322, 1986, pp.808-811.
- [Suc.1] Suck, S.H., Wetmore, A.E., Chen, T.S., and Kassner, J.L., Jr., "Role of Various Water Clusters in IR Absorption in the 8-14- μm Window Region," *Appl. Opt.*, vol.21, 1982, pp.1610-1614.
- [Tho.1] Thomas, M.E., and Nordstrom, R.J., "Line Shape Model for Describing Infrared-Absorption by Water Vapor," *Appl. Opt.*, vol.24, 1985, pp.3526-3530.
- [Two.1] Twomey, S., **Introduction to the mathematics of inversion in remote sensing and indirect measurements**. In *Developments in Geomathematics 3*, New York : Elsevier Scientific Pub. Co., 1977.
- [Vau.1] Vaughan, J.M., **The Fabry-Perot Interferometer: History, Theory, Practice and Applications**. England: IOP, 1989.

- [Vol.1] Volz, A., and Kley, D., "Evaluation of the Montsouris series of ozone measurements made in the nineteenth century," *Nature*, vol.332, 1988, pp.240-242.
- [Wan.1] Wang, J., **An Investigation on The Multiorder Fabry-Perot Interferometer as a Satellite-borne High Resolution Atmospheric Sounder**. Ph.D. thesis, Univ. of Michigan, 1990.
- [Wat.1] Watson, R.T. and Ozone Trends Panel, Prather, M.J. and Ad Hoc Theory Panel, and Kurylo, M.J. and NASA Panel for Data Evaluation, **Present State of Knowledge of the Upper Atmosphere 1988: An Assessment Report**. NASA Reference Publication 1208, 1988.
- [Wie.1] Weinreb, M.P., and Neuendorffer, A.C., "Method to Apply Homogeneous-Path Transmittance Models to Inhomogeneous Atmospheres," *J. Atmos. Sci.*, vol.30, 1973, pp.662-666.
- [Wil.1] Wilksch, P.A., "Instrument Function of the Fabry-Perot Spectrometer," *Appl. Opt.*, vol.24, 1985, pp.1502-1511.
- [WMO.1] **Scientific Assessment of Ozone Depletion: 1991**. World Meteorological Organization, Global Ozone Research and Monitoring Project, report no. 25, 1991.
- [Wya.1] Wyatt, C.L., **Electro-Optical System Design for Information processing**. New York: McGraw-Hill, 1991.
- [Zha.1] Zhao, W., **Thermal Infrared Radiation Transfer in the Planetary Boundary Layer**. Ph.D. thesis, Univ. of Michigan, 1992.